THE BULLETIN OF SYMBOLIC LOGIC
Volume 6, Number 4, Dec. 2000

REVIEWS

The Association for Symbolic Logic reviews selected books and articles in the field of symbolic logic. The reviews were published in ***The journal of symbolic logic*** from the founding of the JOURNAL in 1936 until the end of 1999. The Association moved the reviews to this BULLETIN, beginning in 2000.

The Reviews Section is edited by Herbert Enderton (Coordinating Editor), Thomas Jech, Penelope Maddy, Wolfram Pohlers, and Philip Scowcroft. Authors and publishers are requested to send, for review, copies of books to *Herbert Enderton, Association for Symbolic Logic, Mathematical Sciences Bldg. 7332, UCLA, Los Angeles, California 90095-1566, USA*.

In a review, a reference "JSL XLIII 148," for example, refers either to the publication reviewed on page 148 of volume 43 of the JOURNAL, or to the review itself (which contains full bibliographical information for the reviewed publication). Analogously, a reference "BSL VI 94" refers to the review beginning on page 94 in volume 6 of this BULLETIN, or to the publication there reviewed. "JSL LV 347" refers to one of the reviews or one of the publications reviewed or listed on page 347 of volume 55 of the JOURNAL, with reliance on the context to show which one is meant. The reference "JSL LIII 318(3)" is to the third item on page 318 of volume 53 of the JOURNAL, that is, to van Heijenoort's *Frege and vagueness*, and "JSL LX 684(8)" refers to the eighth item on page 684 of volume 60 of the JOURNAL, that is, to Tarski's *Truth and proof*.

References such as 49***5*** or 280***1*** are to entries so numbered in *A bibliography of symbolic logic* (the JOURNAL, vol. 1, pp. 121–218). Similar references containing the fraction $\frac{1}{2}$ or a decimal point (such as $70\frac{1}{2}$***1*** or 382***7.1***) are to *Additions and corrections to A bibliography of symbolic logic* (the JOURNAL, vol. 3, pp. 178–212).

Articles in ***Handbook of proof theory***

SAM BUSS. *Preface*. ***Handbook of proof theory***, edited by Samuel R. Buss, Studies in logic and the foundations of mathematics, vol. 137, Elsevier, Amsterdam etc. 1998, p. v.

Buss writes in this editor's *Preface*: "Proof theory is the study of proofs as formal objects and is concerned with a broad range of related topics.... This volume contains articles covering a broad spectrum of proof theory, with an emphasis on its mathematical aspects."

It seems to me that the editor's emphasis means non-philosophical or non-foundational aspects. In 1974, Kreisel and Takeuti (JSL L 244) said, "it would be simply absurd to judge proof-theoretic work mainly, let alone solely, by its relevance to Hilbert's original programme, because this programme is itself of dubious relevance." Now is the time to release proof theory from this yoke. Then it turns out that the theory is a collection of results and methods of mathematical interest as collected in the ***Handbook***.

The ***Handbook*** is most welcome in the logic community. Prompt publication of handbooks of other branches in logic is anticipated. I recommend the ***Handbook*** to researchers and graduate students in logic, mathematics, computer science, philosophy, linguistics, artificial intelligence, automated reasoning, and cognitive sciences. This list reflects recent wide applications of proof theory.

The ***Handbook*** contains name and subject indexes but no symbol index. The editor arranges central topics in proof theory carefully and thoughtfully. But publication of a

1079-8986/00/0604-0003/$5.20

handbook of a field results inevitably in depicting a map of the realm. What is missing here is, for example, reverse mathematics by H. Friedman and S. G. Simpson, J.-Y. Girard's Π^1_2-logic and linear logic, Dov Gabbay's labeled deduction, and works done outside the West, e.g. in the former Soviet Union.

The following reviews address the individual chapters in the ***Handbook***. In some cases, a list of typographical errors and questionable points that may puzzle the readers is appended. Many errors that can be easily resolved from the context are not indicated in the lists.

TOSHIYASU ARAI

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

SAMUEL R. BUSS. *An introduction to proof theory*. Ibid, pp. 1–78.

Buss's Chapter I is a good introduction to proof theory. I recommend this exposition to all beginners in the subject.

Proof theory chiefly concerns formal proofs. It is necessary for a beginner to learn and to become familiar with details of proofs as objects to be investigated. This introduction is one of the best training schools for this. Some proofs of theorems are either sketchy or treat a typical case. The reader is encouraged to fill in the remaining cases. I am sure that the task is worthwhile for a beginner.

Mainly sequent calculi and resolution refutations are expounded. It is fair to say that these two formulations are and have been dominant ones in proof-theoretic studies and automated reasoning. Hilbert-style formulations, natural deductions, and non-classical logics such as intuitionistic, modal, and substructural logics (e.g. linear logics) are only briefly mentioned or not included.

In Section 1, propositional sequent calculus, named as PK, is defined and the completeness of the cut-free part of PK is proved by a proof-search method. This in turn implies a cut-elimination theorem for PK. No proof of a numerical (exponential) bound on cut-elimination in propositional logic is given; only a hint is indicated. Although the latter is a folklore result, few textbooks contain its proof. It would be better to include a perspicuous proof by the author. Propositional resolution is explained and its completeness is proved. Then several strategies for implementation of resolution are mentioned.

In Section 2, the first-order sequent calculus LK is defined and the completeness of the cut-free part of LK is again proved by a proof-search method. Furthermore, a clever algorithm for eliminating cut inferences by transforming a given proof is described. This yields a superexponential numerical bound on lengths of resulting cut-free proof from the length of the given proof. Then a so-called free-cut elimination theorem is proved by modifying the algorithm. This means, roughly, that we can eliminate cut inferences whose cut formulas are not non-logical axioms. Hence the theorem has applications for applied logics, i.e., provability in a first-order theory, such as a fragment of arithmetic with induction axioms.

Next as applications of cut-elimination and proof transformations, Herbrand's theorem and Craig's interpolation theorem are shown. Here Herbrand's theorem concerns not only purely existential consequences but also any formulas. The theorem is close to the original theorem due to Herbrand, but a proof is perspicuous and is related to Gentzen's proof of the *Mittelsequenz* theorem. The generalized version of Herbrand's theorem appeared first in the author's paper *On Herbrand's theorem* (***Logic and computational complexity***, Lecture notes in computer science, vol. 960, Springer, 1995, pp. 195–209) and is worth studying even for specialists in proof theory.

Craig's interpolation theorem is proved by splitting sequents, a method due to S. Maehara. Although Skolem's theorem is mentioned, there is no proof-theoretic proof of it. ε-calculi and ε-theorems (Hilbert–Bernays) and Hilbert–Ackermann's ε-substitution method are omitted.

Finally resolution for first-order logic is expounded by using unification.

Errors. Page 17, line l: For 'propositional formulas,' read 'propositional variables.' Page 21, line 14: For 'non-logical,' read 'logical.' Page 33, line 13: Is the language assumed to contain an individual constant? Page 43, line 8 from below: I do not understand the sentence 'This … causes B to be *only weakly introduced in the new subproof*.' Page 65, line 1, (2): $\vee$: *right* in the intuitionistic logic LJ has no side formulas Δ in succedents, as stated in (3).

TOSHIYASU ARAI

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

SAMUEL R. BUSS. *First-order proof theory of arithmetic*. Ibid., pp. 79–147.

In this Chapter II, mainly bounded arithmetics are treated and traditional fragments of arithmetic such as $I\Sigma_n$ are briefly discussed. The results mentioned are divided broadly into two categories:

The first one is the (intensional) definability of, for example, the graph of the exponential function, counting of ones in the binary representation of a number in weak fragments such as $I\Delta_0$. This topic is quite clearly written so that readers can understand the ideas. By using this, (intensional) definitions of sequence encoding and decoding and hence arithmetization of metamathematics (syntax) are obtained. Then Gödel's second incompleteness and its strengthening due to Wilkie and Paris for any extension of S^1_2 are explained. In the latter, avoiding previously used model-theoretic means, the author effectively utilizes (a strong form of) Herbrand's theorem and numerical increases from cut-elimination in Chapter I.

The second one is relationships among fragments of arithmetics. After summaries of easy results, such as $I\Sigma_n = I\Pi_n$, some nontrivial results such as Σ_1-definable functions in $I\Sigma_1$ are primitive recursive functions (Parsons, Mints, Takeuti, independently), Σ^b_i-definable functions in S^i_2 (Buss), and Π_{n+2}-conservation of $B\Sigma_{n+1}$ over $I\Sigma_n$ (Parsons) are shown through free-cut-eliminations of Chapter I and a witnessing technique due to Buss. These are viewed alternatively from Herbrand analyses due to Sieg and model-theoretic proofs (H. Friedman, Paris, Kirby, Wilkie, et al.).

The chapter is recommended as an overview, especially the applications of proof theory to weak fragments. Some results are mentioned without proofs or references. There are unfortunately many errors.

Page 91, line 14: For 'the the,' read 'the.' Similar errors are found in page 92, line 20; page 93, line 16 from below (f.b.); page 121, line 20, and so on. Page 98, line 5 f.b.: Inductive cut is defined in Chapter VIII. Page 113, line 4 f.b.: In connection with Gödel's second incompleteness theorem, it is stressed that the provability predicate should be intensionally defined, i.e., required to obey some derivability conditions, such as demonstrably closed under modus ponens. As shown by Jeroslow it is not necessary that provability predicates enjoy this closure condition in order to have the theorem. In fact the author mentions free-cut free consistency in page 138, line 8, which is based on a non-intensional provability predicate in this sense. Page 115, line 3 f.b. and page 117, line 13 f.b.: In the former place, S^i_2, for example, is declared as a schematic theory, while in the latter place it states that the theory is not schematic. Page 122, line 1: For 'implies,' read 'is implied by.' Page 126: For total functions of $I\Sigma_n$, the references are insufficient and inaccurate. Girard [1987] contains no description of Gentzen's finitary approach to this topic. Ketonen and Solovay [1981] has nothing with model-theoretic constructions. They placed rapidly growing Ramsey functions (Paris–Harrington) in the Grzegorczyk–Wainer hierarchy. It should be mentioned that Kreisel first stressed the topic. Page 132, line 9 f.b.: There is no definition of the complexity class $\Sigma^p_{i+1}/poly$. Page 137, line 7 f.b.: S and T should be interchanged. Page 138, line 9: 'bounded' should be inserted before 'formula' in order to have $BdCon(S) \rightarrow FCFCon(S)$. Page 141, Lemma 4.3.5: I am skeptical that we can pick such a bounding term t. Actually no bound is needed in the following proofs. Page 142, proof of Theorem 4.3.9: Here twice one needs the fact that $S_2 + \exp$ proves a reflection principle saying that if $\phi \equiv \forall x\theta(x)$ is a true Π_1-sentence, then any (bounded) endformula in a bounded proof in $S_2 + \phi$ is true, where $S_2 + \phi$ has extra initial sequents $\rightarrow \theta(t)$ for any term t. Page 143, line 11: There is no definition of S^i_k for $k \neq 2$.

Toshiyasu Arai

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

Matt Fairtlough and Stanley S. Wainer. *Hierarchies of provably recursive functions*. Ibid., pp. 149–207.

This Chapter III is an excellent exposition of provably recursive functions in arithmetic and hierarchies of subrecursive functions. I recommend this for all beginners.

Section 1 motivates the subject well. Classifying provably recursive functions, viz. provably terminating algorithms in a theory containing a weak fragment of arithmetic, is one of the most successful applications of proof theory, and is a source for independence results. This amounts to extracting a bound for the existential quantifier $\exists x$ in a provable Σ^0_1 sentence $\exists x B(x)$ from its derivation in a given theory.

The first such a result was obtained by G. Kreisel using ideas from Ackermann for first-order arithmetic **PA**: Provably recursive functions in **PA** are exactly the functions defined by transfinite recursions on initial segments of ε_0. The class was then organized by Schwichtenberg and Wainer, et al., into extended Grzegorczyk classes $\mathcal{E}^\alpha$ so that the class is stratified into a perspicuous hierarchy. The hierarchy $\{\mathcal{E}^\alpha\}$ is generated from some functions indexed by ordinals and then closed under relative elementary recursiveness.

Traditionally hierarchy generating functions such as the Hardy functions H_α or the fast-growing F_α are defined as transfinite iterations of, for example, the successor function together with diagonalization at the limit stages λ. This requires an assignment of fundamental sequences $\{\lambda_n\}$ converging to λ. (Recently Buchholz, Cichon, and Weiermann found a new mode of generating function, a norm-bounding recursion. This does not require fundamental sequences, cf. Chapter IV.) To get a proper hierarchy, it is commonly required that the function F_β grows faster than (*majorizes*) F_α if $\alpha < \beta$: $F_\alpha(n) < F_\beta(n)$ for almost all n. For this, fundamental sequences have to enjoy a mild condition.

Ordinals in proof theory are not mere set-theoretic objects. These are in fact a term structure having some operations on them. Empirically a term structure for larger ordinals is more complex than one for smaller ones. Therefore it becomes harder to ensure the mild condition for larger structures.

To get rid of this cumbersome verification and adhere to a specific term structure, the authors choose abstract tree-ordinals as representations of ordinals. Then tree-ordinals enjoying the mild condition are singled out to be *structured*, and ordinals indexing hierarchy generating functions are restricted to these. After that, comparisons are established between the functions H_α, F_α, B_α and their honesty, i.e., their computational complexities are bounded by some iterate of themselves. The last entry $B_\alpha = H_{2^\alpha}$ is a refinement of $F_\alpha = H_{\omega^\alpha}$, and reflects faithfully cut-elimination procedure: $B_{\alpha+1} = B_\alpha \circ B_\alpha$, composition corresponding to a cut inference below.

Sections 3 and 4 form the main body of this chapter. The Π^0_2-consequences in **PA** and its fragments are analyzed, based on the method due to Buchholz, presented in an improved version due to Fairtlough. First one embeds finite derivations in infinite ones with the ω-rule. Furthermore the lengths of infinite derivations are assigned so that not only is the conclusion's ordinal λ larger than the premisses' ordinals but also each of the latter ordinals is in essence the nth term λ_n of a fundamental sequence $\{\lambda_n\}$ for λ, where n denotes an input context, roughly speaking, expressing the fact that the number n is substituted for a variable in the embedding. Note that the lengths of infinite derivations depend on the number n

since the induction axiom is replaced by a series of cut rules, and in fact this is the only source of the dependence. Now as usual cuts are eliminated. The length of the resulting derivation is bounded by a finite exponent, say with base 2, of the original one. Finally a bound $B_\alpha(n)$ for $\exists x C(n,x)$ is extracted from the cut-free derivation of length α since a cut-free derivation of a Σ^0_1 sentence is essentially finite, and the ordinal assignment is designed so that the number $B_\alpha(n)$ increases when we go down to the conclusion of an inference. This elegant proof is due to Buchholz, and extracts a hierarchy generating function B_α as a bound for Π^0_2-consequences directly from its derivation, i.e., it throws away the intermediate step, viz. first extracting a bounding function defined by a transfinite recursion and then showing such a function is bounded by a B_α.

Unfortunately the authors' presentation of the Bounding Lemma 3.15 on page 170 is liable to cause misunderstanding. Namely in the statement of Lemma 3.15.2, the second sentence should read something like this (according to authors' response): "Suppose also that n bounds the value of every (closed) quantifier-bound occurring in $\mathbf{A}$, *and* that whenever a cut occurs in the derivation with premisses $\mathbf{A}, \neg C$ and $\mathbf{A}, C$ with context $m : \mathsf{N}$, then m similarly bounds the value of every closed quantifier-bound in $\mathbf{A}, C$." Further, Definitions 3.13 should define "$n : \mathsf{N}$ is true at m" to mean that $n \leqq m$,

In the final sections Girard's hierarchy comparison theorem is well expounded without the heavy Π^1_2-logic. This theorem establishes how far we go forward as slow-growing functions catch up with fast-growing ones. Finally theories with transfinite induction are analyzed with respect to Π^0_2-consequences. The analysis of **PA** is easily extended to such theories, cf. the following chapter by Pohlers.

TOSHIYASU ARAI

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

WOLFRAM POHLERS. *Subsystems of set theory and second-order number theory*. Ibid., pp. 209–335.

In Chapter IV, Pohlers presents a sample of *ordinal analyses* of subsystems of set theory and second-order number theory by Gentzen-type methods, i.e., analyses based on cut-eliminations of (mainly infinite) derivations. Analyzed theories are subsystems of a set theory **KPi** axiomatizing recursively inaccessible universes. This is the longest chapter in the ***Handbook***, reflecting the status of ordinal analyses in proof theory. Ordinal analysis by Gentzen-type methods is one of the technically most advanced and complicated topics, and has been a central theme in proof theory.

The chapter is the first exposition of the topic for fairly strong theories to appear in book form. It contains many reflections on results and it explains motivations (heuristics) behind complicated techniques.

An ordinal analysis of an axiomatic theory **Ax** including first-order number theory **PA** is a computation or description of the *proof-theoretic ordinal* $\|\mathbf{Ax}\|$ of **Ax**, i.e., the limit of ordinals α such that **Ax** proves the fact that $\prec_\alpha$ is a well ordering for some recursive relation $\prec_\alpha$ on ω of type α. For a recursive and sound theory **Ax**, $\|\mathbf{Ax}\|$ is recursive. Ordinals in ordinal analyses are represented by recursive term structures having operations and relations. Analyses so far obtained yield, at the same time, constructive or computational content hidden in derivations in **Ax** as exemplified in the chapter. An ordinal analysis of **Ax** includes two aspects. First one devises a recursive relation $<_{\mathbf{Ax}}$ with a term structure by analyzing (infinite representations of) derivations in **Ax**, and shows that the order type of $<_{\mathbf{Ax}}$ gives an upper bound for $\|\mathbf{Ax}\|$. Secondly one shows in **Ax** that each initial segment of $<_{\mathbf{Ax}}$ is well founded. In this long chapter only the first aspect is expounded. For this aspect, i.e., an upper bound result, it suffices to show that any Π^1_1-consequence A in **Ax** has a cut-free ω-derivation of height $<_{\mathbf{Ax}} \alpha_A$ for some ordinal (integer) α_A. Here a cut-elimination is involved.

Now let us make an excursion on the untrodden region following Pohlers. The chapter consists of three sections.

In the first section the proof-theoretic ordinal $\|\mathbf{Ax}\|$ of a set theory $\mathbf{Ax}$ is reformulated as a minimum Π^Ω_2-model of $\mathbf{Ax}$ ($\Omega = \omega_1^{\mathrm{CK}}$): Suppose that the constructible universe L is a model of a recursive theory $\mathbf{Ax}$. Then the least ordinal α such that $L_\alpha \models A$ for any Π^Ω_2-consequence A^{L^Ω} of $\mathbf{Ax}$, i.e., $\mathbf{Ax} \vdash A^{L^\Omega}$ is called the Π^Ω_2-*ordinal* of $\mathbf{Ax}$ and denoted $\|\mathbf{Ax}\|_{\Pi^\Omega_2}$. Under a mild condition on $\mathbf{Ax}$ it is shown that $\|\mathbf{Ax}\|_{\Pi^\Omega_2} = \|\mathbf{Ax}\|_{\Sigma^\Omega_1} = \|\mathbf{Ax}\|$ (the middle ordinal is defined analogously with Σ^Ω_1-consequences). Thus the proof-theoretic ordinal $\|\mathbf{Ax}\|$ receives a clear set-theoretic meaning. Furthermore the ordinal coincides with the supremum of the *truth complexities* $\mathsf{tc}(\forall XF(X))$ of Π^1_1-consequences $\forall XF(X)$ of $\mathbf{Ax}$, where $\mathsf{tc}(\forall XF(X))$ is defined to be the height of the shortest cut-free ω-derivation of $F(X)$ if such a derivation exists. Here the main technical lemma is stated as Boundedness Lemma 1.3.8 by which we can extract an upper bound of the order type of a relation $\prec$ from the height of a cut-free derivation proving the well-foundedness of $\prec$. The lemma goes back to K. Schütte. An improved version due to A. Beckmann is fully explained.

In the second section a conservative extension $\mathbf{NT}$ of $\mathbf{PA}$ is analyzed as a paradigmatic example. The main lemma here is the Elimination Lemma 2.1.2.8 stating that we can lower the complexity of cut formulas by 1 at the cost of an exponential increase of the height of derivation. For finite derivations the lemma is stated in the Chapter I, page 41. From the lemma we see that the height of a cut-free derivation is bounded by the ρth iterate of the exponential function (i.e., Veblen function $\varphi\rho\alpha$ defined in Section 1) of height α of the original derivation, when complexity of cut formulas in the original derivation is bounded by ρ. This leads to ordinal analyses of predicative fragments of second-order number theory: Gentzen's theorem $\|\mathbf{NT}\| = \varepsilon_0$, and the Feferman–Schütte identification of the limit of predicatively secured ordinals with the first strongly critical ordinal Γ_0, the least non-zero ordinal closed under the (binary) Veblen function.

Next the theme turns to Π^0_2-consequences in $\mathbf{NT}$: provably recursive functions in $\mathbf{NT}$ as treated in Chapter III. Methods here are due to A. Weiermann and are based on a finite miniaturization of techniques in impredicative proof theory, namely local predicativity method due to Pohlers and its "simplified" version (operator-controlled derivations) due to W. Buchholz. The latter technique is explained in Section 3. Very roughly speaking, the finite miniaturization is obtained by replacing the ordinal α by the complexity (called the *norm*) of its representation, for example, the number of symbols occurring in the term α. Then provably recursive functions in $\mathbf{NT}$ are classified in terms of functions $\psi_\omega(\alpha)$ defined by norm-bounded transfinite recursion, as opposed to the approach in Chapter III based on fundamental sequences. The function ψ_ω collapses a countable ordinal α to a finite one $\psi_\omega(\alpha)$.

Section 3 is the main body of this contribution, and treats impredicative systems. First impredicative axiom systems for number theory are defined and elementary facts about them are established. Theories of positive elementary inductive definitions on ω and of transfinitely iterated ones are defined. Then these theories are related to iterated Π^1_1-comprehension axioms. This result is due to S. Feferman and is based on the tree analysis of Π^1_1-predicates due to Lusin–Sierpiński and Kleene. Secondly these systems for number theory are embedded into set theories. In a universe containing an increasing sequence of admissible sets one can iterate inductive definitions or equivalently Π^1_1-comprehension axioms.

In the final Subsection 3.4, ordinal analyses of subsystems of $\mathbf{KPi}$ are fully expounded. The method is based on the *operator-controlled derivations*, a highly sophisticated method due to Buchholz. Let I be the first recursively inaccessible ordinal, and $\kappa \leqq I$ a Π_2-reflecting ordinal. In eliminating cut inferences one of whose cut formulas is a conclusion $\exists z \in L_\kappa A^z$ of an axiom for L_κ, one has to extract a witnessing set L_η ($\eta < \kappa$) for z from a derivation of the premiss A^{L_κ} of the axiom. This involves collapsing operations on infinite derivations by which we can collapse an infinite derivation of a $\Pi_2^{L_\kappa}$-endformula with the height $\geqq \kappa$ down to one of height $< \kappa$, together with *collapsing functions* $\psi_\kappa(\alpha) < \kappa$ by which we can describe heights of collapsed derivations. Operators $\mathcal{H}(X)$ on sets of ordinals X, new data equipped with derivations, display a close relationship between the former operations on derivations and the latter functions on ordinals succinctly. The set $\mathcal{H}_\gamma(X)$ of ordinals is defined with an ordinal index γ to be a Skolem hull of $X \cup \{0, I\}$ under the functions $+$, $\lambda\alpha.\omega^\alpha$, φ, $\lambda\alpha.\omega_\alpha$, $\lambda\alpha.\alpha^+$ (the least initial number greater than α), and their inverses together with the collapsing functions $\lambda\kappa\lambda\xi < \gamma.\psi_\kappa(\xi)$. The latter is defined by $\psi_\kappa(\gamma) = \min\{\beta : \kappa \in H_\gamma(\beta), H_\gamma(\beta) \cap \kappa \subseteq \beta\}$ where $\beta = \{\alpha \in \mathrm{On} : \alpha < \beta\}$. $\mathcal{H}_\gamma$ and $\psi_\kappa(\gamma)$ are defined simultaneously by induction on γ. As a sample result $\|\mathbf{KPi}\| \leqq \psi_\Omega(\varepsilon_{I+1})$ is shown. In fact equality holds since G. Jäger showed $\psi_\Omega(\varepsilon_{I+1}) \leqq \|T_0\| \leqq \|\mathbf{KPi}\|$ for Feferman's constructive theory T_0. The chapter concludes with a table summing up results for subsystems of **KPi** due to Pohlers, Buchholz, Jäger, Rathjen, et al.

Errors. Page 216, page 218, page 228: Soundness of the axiom system **Ax** is assumed, and the axiom $V = L$ of constructibility is to be a consequence of the set theory **Ax**, or a similar assumption seems to be necessary. Page 216, line 18: Any axiom *except* the foundation scheme is at most Π_3. Page 218, line 4: What does the sentence '$\mathbf{L}_\kappa$ is definable in **Ax**' mean? In the chapter only the case $\kappa = \omega_1^{\mathrm{CK}}$ is needed. Page 219, line 4 from below (f.b.): There is no primitive recursive function computing values of closed terms in a language with all primitive recursive functions. Page 238: This is essentially the same proof given in Chapter III. Page 262, line 9: Insert N after P. Page 263, line 5 f.b.: The proof predicate $Proof_\infty$ as described in the bottom of the page seems to be Π_1^0. Page 266, line 8: In the premiss, $Proof_{\mathbf{Ax}}$ should be $Proof_{\mathbf{NT}_{\prec|}}$. Page 266, line 12: This may be *infinite* because of the presence of replication rules. Finiteness follows from transfinite induction. Page 271, line 10: According to J. Zucker (*Iterated inductive definitions, trees and ordinals*, ***Metamathematical investigation of intuitionistic arithmetic and analysis***, Lecture notes in mathematics, vol. 344, Springer-Verlag, 1973, pp. 392–453), it was G. Kreisel who introduced theories of finitely iterated inductive definitions in a review, ***Zentralblatt für Mathematik und ihre Grenzgebiete***, vol. 106 (1964), pp. 237–238. Page 275 line 5 f.b.: The formula A contains a parameter X. The proof works only for **Aut-BID** since, for example, in $\mathbf{BID}_\nu^2$ no parameter may occur, cf. page 274, line 6. Page 217, line 22, page 283, line 4 f.b., page 291, line 4 f.b.: ω is an admissible ordinal in the first place but not in the latter ones. Page 288, line 11 f.b.: Uniqueness of the function f follows only if, for example, f has to be an onto map. Page 313, line 15, line 3 f.b.: (199) should be (181).

TOSHIYASU ARAI

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

JEREMY AVIGAD and SOLOMON FEFERMAN. *Gödel's functional ("Dialectica") interpretation*. Ibid., pp. 337–405.

Chapter V, by J. Avigad and S. Feferman, is a good, updated survey on Gödel's functional (Dialectica) interpretation (*D-interpretation* for short) and its extensions and refinements, including recent developments.

D-interpretation reduces an intuitionistic theory to a functional (quantifier-free) theory as follows. First one translates a formula φ to a formula $\varphi^{\mathrm{D}} \equiv \exists x \forall y \varphi_{\mathrm{D}}$ where φ_{D} is a quantifier-free formula in a higher-order language. Then one shows that there exists a functional $f(y)$ such that $\varphi_{\mathrm{D}}(f(y), y)$ by induction on the derivation of φ. Various forms of induction are replaced here by suitably related forms of recursion, for example, complete induction is replaced by primitive recursion of finite types. Thus one can extract directly a constructive-computational content of the intuitionistically derivable formula φ from its derivation. A disadvantage lies in the fact that the complexity of the interpretation grows quickly with the formulas, and becomes unmanageable in practice.

Beginning with an excellent overview and a list of works on the interpretations, the chapter presents D-interpretation of first-order and higher-order arithmetic using primitive recursive functionals of finite type. It shows the equivalence $\varphi \leftrightarrow \varphi^{\mathrm{D}}$ assuming the axiom of choice (AC), independence of premiss (IP′), and (a generalization of) Markov's principle (MP′). It is argued that the first one is accepted in constructivism while the latter two are questionable. When combined with double-negation translation, D-interpretation yields a reduction of a classical theory to a functional theory, *ND-interpretation* φ^{ND}. In classical logic, the axiom schema (AC) is proof-theoretically strong: in fact it yields the full comprehension axiom of finite types. The equivalence $\varphi \leftrightarrow \varphi^{\mathrm{ND}}$ is shown under the assumption of a weak form (QF-AC) of (AC) applied only to quantifier-free formulas. Some consequences of the interpretation follow, for example, the fact that every provably recursive function is denoted by a primitive recursive functional.

Section 4 takes up the issue of models of primitive recursive functionals: hereditarily recursive operations HRO as a model of intensional equality of higher types, hereditarily effective operations HEO as a model of extensional equality, and the term model as a computational and syntactical model. The last one is defined through rewrite rules and requires showing the normalization theorem. The theorem is proved first through reducibility (computability) predicates and through infinitely long terms, both methods due to W. Tait.

After a short section on interpretations of fragments of arithmetic including bounded arithmetic of Cook and Urquhart, Spector's interpretation of second-order arithmetic with the unrestricted comprehension axiom is presented in the form of W. Howard: an interpretation of bar induction by bar recursive functionals. In the following three sections it is demonstrated that D-interpretation can be used to obtain information regarding a number of interesting subsystems of second-order arithmetic. Namely, D-interpretations are applied to problems on the subsystem WKL_0 with weak König's lemma, subsystems such as Σ^1_i-DC, $(i = 1, 2)$ with restricted dependent choices, and theories of tree ordinals. Thereby we reobtain known results: (1) H. Friedman's theorem stating that WKL_0 is Π^0_2-conservative over PRA, a proof due to U. Kohlenbach via the hereditary majorizability by Howard. (2) By using non-constructive principles, i.e., the minimum operator and the Kleene basis operator, some subsystems with a restricted choice axiom are reduced to iterated comprehension axioms, a proof due to Feferman of H. Friedman's result. (3) A classical theory of tree ordinals is ND-interpreted in a functional theory, an unpublished work due to Feferman and related to Howard's work.

In the last section, functional interpretations based on polymorphism are presented: J.-Y. Girard's interpretation of the comprehension axiom by means of second-order polymorphic λ calculus F and Avigad's one of classical theories $\widehat{ID}_n$ of finitely iterated fixed points in predicative polymorphic functionals P_n.

Errors. Page 385, line 9: Π^1_2 (in Π^1_2-AC) should be Σ^1_2 or Π^1_1. Page 388, line 7: In the standard terminology in constructivism (e.g. page 66 and page 232 in Troelstra and van Dalen, ***Constructivism in mathematics***, Volume I, North-Holland, 1988) strictly positive occurrences are wider than strongly positive ones. Here is meant the former.

Toshiyasu Arai

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

A. S. Troelstra. *Realizability*. Ibid., pp. 407–473.

Realizability was introduced by S. C. Kleene as a semantics for intuitionistic arithmetic. Since then, many notions of realizability have been defined for various theories and applications. Chapter VI, *Realizability*, by A. S. Troelstra provides a good survey of the material. The presentation is fairly standard. The chapter contains suggestions for further reading, historical comments, and many references.

Section 1 is devoted to a full exposition of the earliest and most basic notion of realizability, Kleene's realizability by numbers. To each arithmetical formula A, a formula $x \mathsf{rn} A$ (read "x numerically realizes A") is associated so that the realizer x hereditarily provides witnesses for the constructive truth of existential quantifiers and disjunctions, and in implications carries this type of information from premiss to conclusion by means of partial recursive operations. By introducing a conservative extension $\mathbf{HA}^*$ of $\mathbf{HA}$ based on logic of partial terms, partial recursive function applications in defining $x \mathsf{rn} A$ are written succinctly. Some standard results on rn-realizability such as soundness, $\exists$-free formulas as self-realizing ones, and characterization by extended Church's thesis (ECT_0) are proved. In order to establish, for example, explicit definability for numbers (EDN), a combination of realizability with truth, $x \mathsf{rnt} A$, is defined by declaring $x \mathsf{rnt}$-realizes an implication $A \rightarrow B$ only if it is true.

Section 2 defines abstract realizability and function realizability. The former is defined through the theory **APP** whose models are partial combinatory algebras. In the latter, rf-realizability, the realizing objects are (number-theoretic) functions, and applications are partial continuous applications. Namely each function α is regarded as encoding partial continuous functionals $\alpha(\beta) = x$ from Baire space to the set of natural numbers or functionals $\alpha \mid \beta = \gamma$ on Baire space. Then corresponding results on this realizability are shown for intuitionistic elementary analysis **EL**.

In Section 3, modified realizability, a realizability for intuitionistic finite-type arithmetic(s) $\mathbf{HA}^\omega$ is defined. Here a realizing object is in a finite-type structure with application. By using the realizability, it is shown, for example, that $\mathbf{HA}^\omega$ is closed under the rule of choice ACR. In the next section it is shown that how closure under the rule ACR may be combined with the notion of majorizable functional to obtain closure under the fan rule, an intuitionistic analogue to König's infinity lemma.

The following two sections briefly discuss Lifschitz realizability and extensional realizability. The former is applied to show Church's thesis with uniqueness ($\mathrm{CT}_0!$) does not imply Church's thesis CT_0 in **HA**. The latter combines realizability with a relation, "two objects equally realize a formula."

Sections 7 and 8 discuss realizability for intuitionistic second-order **HAS** and higher-order arithmetic **HAH** with full comprehension axioms. Realizability for higher-order arithmetic is defined as a truth-value semantics. Let Ω denote the set of truth values (e.g. the power set of the set of natural numbers). To each formula in a higher-order language, an element of Ω—the truth value of the formula—is associated. Then the truth value is regarded as the set of realizers of the formula, i.e., the formula is realizable (called *valid*) iff the value in Ω is inhabited (a non-empty set). An interesting corollary is that **HAS** is consistent with the uniform principle UP stating that you can exchange the order of a universal second-order quantifier for sets and an existential first-order quantifier. The modeling of higher-order logic is explained in terms of a category Eff, in fact it is shown to be a topos. Finally for **HAS** it is shown that the category-theoretic model is equivalent to a concrete rn-realizability in which $x \mathsf{rn} XT$ is defined as $X^*(x, t)$ with a new set variable X^*, representing the realizability predicate.

The final section summarizes further works such as realizability for set theory and for subsystems of intuitionistic arithmetic, formula-as-types realizability, completeness questions for realizabilities and applications to computer science. Various results and original papers are mentioned.

Errors. Page 426, line 15: For the second '$\phi \downarrow$', read '$t \downarrow$'. Page 433, line 11 from below (f.b.): For 'αx', read 'αy'. Page 436, line 12 f.b.: For 'closure under the fan rule,' read 'closure under the rule of choice.'

TOSHIYASU ARAI

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

Giorgi Japaridze and Dick de Jongh. *The logic of provability*. Ibid., pp. 475–546.

It is fair to say that a mathematical theory of the logic of provability began with Solovay's pioneering work in 1976 (JSL XLVI 661). He showed that propositional reasoning about the formalized provability predicate for arithmetic is decidable and can be adequately described in arithmetic itself. This chapter contains a detailed account of Solovay's work and its extensions, and thereby surveys the current state of affairs of the subject, namely extensive works done by Artëmov, Beklemishev, Japaridze, de Jongh, Lindström, Montagna, Visser, et al. since 1980. Some of papers by the Moscow school were written in Russian and are not easy to access. This chapter remedies the inconvenience and is welcomed by a wide audience.

Solovay's first arithmetic completeness theorem says that a modal formula A in the language having a single modal operator $\Box$ is provable in a propositional modal logic **L** iff all arithmetic realizations A^* are provable in a (hence any) Σ_1-sound r.e. extension T of **PA**. A^* denotes a result of substituting arithmetic sentences for propositional variables, and replacing the modal operator $\Box$ by a *standard* Σ_1-provability predicate Pr_T for T, i.e., a predicate satisfying the three Hilbert–Bernays–Löb–Kreisel conditions, where we assume that there are infinitely many witnesses ("proofs") of the theoremhood Pr_T for every theorem. Although these conditions are fulfilled for the usual provability predicate, Pr_T need not to be a canonical one. The modal logic **L** extends the minimal normal logic **K** by adding the schema $\Box(\Box A \to A) \to \Box A$, a modal counterpart of the formalized Löb's theorem. Löb's third condition $\Box A \to \Box\Box A$ is then derivable in **L**. Solovay's second theorem concerns the truth of A^* instead of provability in T. The modal logic **S** for the logic of truth is obtained from **L** by adding the schema $\Box A \to A$ for the local reflection principle, and is not a normal logic. The proofs of these theorems involves two steps: modal completeness and arithmetic completeness. The former establishes that if $\nvdash_{\mathbf{L}} A$, then there exists a finite Kripke model M for **L** in which A is not valid. The latter, in turn, constructs an arithmetic realization A^* from the countermodel M of A so that, for example, in the first theorem for provability, $T \nvdash A^*$. This is done by simulating finite Kripke models within the arithmetic theory T with the Solovay function, a genuine generalization of Rosser's construction. Sections 2–3 give the detailed account of Solovay's proof.

Sections 4–8 discuss further results in the single modal operator: provability logic or truth provability logic of T (i.e., the set of all modal formulas any arithmetic realizations of which are provable in T or are true, respectively), and arithmetic complete modal logics (i.e., those principles of provability in T that can be verified by means of a theory U). Surprisingly these modal logics can be classified in terms of a few families of formulas, mainly expressing the consistency height of T, i.e., the number of times we can iterate consistency statements for T (work of Visser, Beklemishev).

After brief discussions of bimodal provability logics, Rosser ordering, and logic of proofs (Artëmov), Sections 11–15 deal with interpretability and partial conservativity of r.e. extensions of **PA**, and modal logics of them. In Section 12 authors give a full exposition of theorems of Orey and Hájek saying interpretability is nothing but Π_1-conservativity, of Japaridze saying cointerpretability—a dual of interpretability—is nothing but Σ_1-conservativity, and of Lindström saying faithful interpretability is equivalent to interpretability and cointerpretability. Section 13 contains the modal completeness of a bimodal logic **ILM** for the interpretability (or equivalently the Π_1-conservativity) logic with respect to *finite* Veltman-frames. **ILM** contains a binary connective $A \rhd B$, which is interpreted as $T + B$ is interpretable in $T + A$, or equivalently $T + B$ is Π_1-conservative over $T + A$. Then in Section 14 arithmetic completeness of **ILM** is proved in an informative way. The proof is a refinement of Solovay's, and the exposition is very readable. Section 15 discusses the tolerance and cotolerance logics having the single variable-arity modal operator $\Diamond(A_1, \dots, A_n)$ (read "there are consistent extensions T_i of $T + A_i$ so that T_{i+1} is interpretable in T_i").

Finally Section 16 discusses predicate provability logics. First it is shown that the predicate provability logic is not axiomatizable, a negative result. Then some positive results are proved. Notably we can retain the arithmetic part of Solovay's theorem in the predicate level. A simulation of finite Kripke counter-models within an arithmetic theory is fully expounded. Here 'finite' means that not only the set of worlds W is finite but also domains $\{D_w\}_{w\in W}$ have to be finite.

Errors. Page 483, proof of (e): On page 481, line 11, T is assumed only to be Σ_1-sound, and Lim_r is a Σ_2-sentence. Page 495, line 12, about provable closure under modus ponens, i.e., Löb's second derivability condition on Rosser's provability predicates: Although it is not clear what is meant in this line, there are standard provability predicates for which Rosser's provability predicate enjoys provably the second or third condition, cf. lines 16–21 on page 496. Page 499, line 4 and line 7 from below: In the latter we need to have that if $\mathrm{Prf}(\overline{n}, \ulcorner F^* \urcorner)$ for some n, then $\mathbf{PA} \vdash F^*$, but in the former only the converse is required.

Toshiyasu Arai

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

Pavel Pudlák. *The lengths of proofs*. Ibid., pp. 547–637.

The subject of the lengths of proofs in proof systems has recently attracted attention since the subject is very close to studies of bounded arithmetic and (computational) complexity theory. The central problem in this area is to determine the minimal complexity of a proof of a theorem in a given proof system. In particular, the problem is to ask whether a proof system can have feasible (i.e., manageable) proofs of theorems. Thus it is related to and is motivated by a finitistic question on human or mechanical reasonings when the latter are represented in proofs in a fixed proof system.

The chapter is a good survey of an important subject written in a book form for the first time, and contains expositions of some unpublished results and (Section 11) bibliographical remarks for further reading.

The most important measure of the complexity of proofs is the *size* of a proof, i.e., the total number of symbols occurring in the proof. The next one is the *number of steps* (proof lines). The chapter consists of two parts. The first part, Sections 3–7, mainly concerns first-order proof systems. The second, Sections 8–10, concentrates on propositional proof systems.

In Section 3, Solovay's unpublished construction of short formulas $\{\varphi_n\}$ with short proofs is fully expounded. Here 'short' means that the sizes are bounded by a polynomial in n. The formulas φ_{n+1} should be equivalent to a formula defined from the previous one φ_n in a uniform way. By using the construction, it is shown that some schematically provable theorems have short proofs; for example, short partial truth definitions are obtained so that Tarski's condition for the truth definitions can be proved by polysize (polynomial size) proofs, and shortenings by *cuts*. In Section 6, the author applies the former to obtain an upper bound on sizes of proofs in a theory T of finite consistency statements $\mathrm{Con}_T(n)$ which say that T is consistent up to proofs of size $\leqq n$. A proof similar to the proof of Gödel's second incompleteness theorem gives a lower bound for $\mathrm{Con}_T(n)$, first obtained by H. Friedman in an unpublished note. These bounds are close to each other and are of the form n^c for constants c. Further, a non-elementary *speed-up* of Gödel–Bernays set theory **GB** over Zermelo–Fraenkel set theory **ZF** is obtained in Section 7 by using $\mathrm{Con}_{\mathbf{ZF}}(n)$.

The (first-order) unification algorithm is applied in Section 4 to yield a most general proof with the same *proof skeleton* (a concept due to Parikh and Orevkov) of a given proof. Thereby we can bound the (logical) depths of formulas in terms of complexities of a given proof. The depth of a formula is defined by ignoring terms occurring in it. It is mentioned that bounding the size of a proof in terms of the number of steps and the proved formula is a difficult problem, since (except for cut-free proofs) we do not know how to bound the complexities of terms using these data. Given a number k of steps and a formula φ, and searching for a proof of φ in k steps in, for example, sequent calculus for predicate logic, the unification approach produces second-order unification problems which are known to be undecidable. In fact Buss showed that the number-of-steps problem for the sequent calculus is undecidable. Luckily the technique of unification can be applied to solve affirmatively the famous *Kreisel's conjecture* for finite fragments of arithmetic enjoying a mild condition as explained in Section 4. Besides these applications of unification, the author explains Krajíček's fundamental result on the polynomial equivalence of tree-proofs and sequence-proofs in proof systems having cut inference (modus ponens).

Section 5 proves a result due to Orevkov: non-elementary speed-up of first-order calculus having cut inference over its cut-free part. Sections 8–10 treat the currently most active area: propositional logic and bounded arithmetic.

Frege systems $\mathcal{F}$—schematic systems having finite schematic rules—are shown to be polynomially equivalent (Cook and Reckhow). While Frege systems concern truth of formulas, *extension* (extended) Frege $e\mathcal{F}$ and *substitution* Frege $s\mathcal{F}$ concern validity, i.e., in $e\mathcal{F}$ we can abbreviate long formulas by atoms and in $s\mathcal{F}$ we can substitute any formula for an atom in a proved formula. The fundamental results on these systems are proved in Section 8: $s\mathcal{F}$ is polynomially equivalent to $e\mathcal{F}$. And the number of steps in $\mathcal{F}$, the number of steps in $e\mathcal{F}$, and size in $e\mathcal{F}$ are polynomially related. Furthermore by using a game-theoretic technique, it is shown that the number of steps in tree-$\mathcal{F}$ is polynomially bounded by the number of steps in sequence-$\mathcal{F}$. The section ends with a catalogue of propositional proof systems, such as quantified Frege, *Nullstellensatz*, and cutting plane proof systems.

In Section 9, as an example of lower bounds proofs, a full proof is given of A. Haken's theorem stating that refutation-proofs of propositional pigeonhole principles PHP_n in resolution must be exponentially large. Originally it was proved by *bottleneck counting*; the proof method here due to Ajtai is based on *random restriction*. Using the sophisticated method, one can show the existence of a restriction on atoms through which suitable Boolean algebras could be defined for purported short refutation-proof of PHP_n.

Next the simple but powerful technique for lower bounds based on effective interpolation theorems is explained. The technique assumes the existence of a pair of disjoint NP sets that are supposed not to be separated by a set in P. By using exponential lower bounds for the monotone circuits and a Lyndon-style effective interpolation theorem for resolution, an exponential lower bound for resolution is obtained. Further it is reported that an effective interpolation theorem for Frege systems $\mathcal{F}$ does not hold, assuming a widely accepted conjecture in complexity theory. This result casts a doubt on the applicability of the technique to strong proof systems.

In Section 10, a close relationship between bounded arithmetics and propositional proof systems is explained. A landmark here is the S. Cook's theorem stating that $e\mathcal{F}$ is *the* propositional proof system for bounded arithmetic S^1_2 for polytime computable functions. Namely any Π^b_1 theorem in S^1_2 can be translated to Boolean tautologies having polysize proofs in $e\mathcal{F}$. Conversely the reflection principle for $e\mathcal{F}$ is provable in S^1_2. This yields two corollaries. The first says that $e\mathcal{F}$ is the strongest system that can be verified in S^1_2 to be sound. The second says that S^1_2 does not prove NP $\neq$ co-NP if there is a superpolynomial lower bound for proofs in $e\mathcal{F}$. Finally a link between optimal proof systems and finite consistency statements is mentioned.

Errors. Page 558, line 6 from below (f.b.): The subscript n in $\forall y \beta_n$ should be $n+1$. Page 607, line 12 f.b.: For '$i \in R$', read '$j \in R$'. Page 607, line 6 f.b.: The subscript $i \in R$ of $\bigvee$ should be $i \in D$. Page 607, line 5 f.b.: For '$j \in D$', read '$i \in D$'. Page 608, line 16 f.b.: For '$g' \in V_T$', read '$g' \in V_{T'}$'. Page 610, formula (55): $n(|A| - a)$ is the denominator. Page 611, line 7 f.b.: For '$g \in M_n$', read '$g \in M_{n,n'}$'. Page 613, lines 13–15: The arguments of $\neg\Phi$

should be $(\bar{a}, \bar{q})$. Page 622, line 5: Second-order bounded arithmetics are undefined in the ***Handbook***.

Toshiyasu Arai

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

Gerhard Jäger and Robert F. Stärk. *A proof-theoretic framework for logic programming*. Ibid., pp. 639–682.

Chapter IX by G. Jäger and R. F. Stärk is a survey article of the authors' recent work on a deductive approach to logic programming. Some familiarity with logic programming is presupposed.

For definite logic programs P the situation is quite simple: a match between declarative and procedural interpretations. Namely a ground atom A is a logical consequence of P iff $P \cup \{\leftarrow A\}$ has an SLD-refutation. Further these are equivalent to $M_P \models A$ for the least Herbrand model M_P of P, which is the least fixed point of a continuous (i.e., monotonic and finitary) operator T_P associated with P, and hence is the ωth iterate of T_P starting with the always falsifying model. Once one considers (general) logic programs and (general) goals that contain negative literals in their bodies, one encounters an obscure picture. This is partly because we do not know how to represent and get negative information when positive contents are inductively defined. Proposals to overcome the defect are, on the declarative side, *three-valued* (and four-valued) structures by Fitting, Kunen, et al., and on the procedural side, SLDNF-resolution with Clark's *negation as failure rule*. Several researchers have discussed the completeness of SLDNF-resolution under three-valued or two-valued structures with respect to some classes of logic programs. The authors provide a proof-theoretic framework for the matter by introducing a deductive system $\mathcal{R}(P)$.

Section 2 begins by defining (lower) three-valued structures, where truth values are $\{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ ($\mathbf{u}$ for undefined). These are ordered by the *information ordering* $\sqsubseteq$, i.e., the reflexive closure of $\mathbf{u} \sqsubseteq \mathbf{x}$ for $\mathbf{x} \in \{\mathbf{t}, \mathbf{f}\}$. Truth values of equations $a = b$ are in $\{\mathbf{t}, \mathbf{f}\}$. For structures $\mathcal{M}, \mathcal{N}$ agreeing on everything except interpretations of relations R, $\mathcal{M} \sqsubseteq \mathcal{N}$ iff $\mathcal{N}(R)$ extends $\mathcal{M}(R)$ in the sense of $\sqsubseteq$. An elegant and clever formulation of Clark's equational theory $\mathrm{CET}_\mathcal{L}$ for a language $\mathcal{L}$ is given so that equality axioms are closed under cut rules. A structure is said to be an *equational structure* if the equality axioms in $\mathrm{CET}_\mathcal{L}$ are true in it. A three-valued equational structure is called *adequate to P* if the Clark completion $\mathrm{comp}(P)$ of P is true in it. The *P-envelope* $\mathcal{M}[P]$ is defined as a many-valued analogue to the immediate consequence operator T_P. The operator is monotonic but in general not continuous. Hence the closure ordinal to reach the least fixed point $\mathcal{IM}_P$ starting with an always undefined structure $\mathcal{M}$ may be larger than ω. An example for this phenomenon is given. It is shown that $\mathcal{IM}_P$ is the least three-valued structure that is adequate to P.

After these model-theoretic discussions, calculi $\mathcal{R}(P)$ for a logic program P are described in Section 4. $\mathcal{R}(P)$ is formulated in Tait calculus by augmenting equality axioms in $\mathrm{CET}_\mathcal{L}$ and inference rules for $\mathrm{comp}(P)$. Axioms $\Gamma, \neg A, A$ for an atom A are called *identity axioms*. Then derivations without identity axioms, called *identity-free derivations*, are singled out. It is indicated that, by using Schütte's deductive chains, we see that identity-free (and cut-free) derivability is equivalent to validity (i.e., to receive the truth value $\mathbf{t}$) in all three-valued equational and adequate structures. This is a dual to the fact due to Schütte and Girard that cut-free derivability is equivalent to non-falsehood in three-valued structures. Using the completeness of identity-free derivations together with an observation that iterated formation of envelopes is the semantic counterpart of the program rules, Corollary 4.2.5 concludes that three-valued validity in equational structures and structures adequate to P implies truth in finitely iterated envelopes starting with, for example, an always undefined structure, a result due to Kunen. Kunen's proof is based on an ultrapower construction. Thus by introducing (identity-free part of) calculi $\mathcal{R}(P)$ we get a deductive approach to three-valued validity, and thereby a perspicuous proof of Kunen's result.

Section 5 first defines SLDNF-resolution, following Kunen, by means of signed sets $\mathbf{R}(P)$ and $\mathbf{F}(P)$. It then presents, following Buchholz, a theoretically more transparent approach to SLDNF-resolution by means of sets $\mathbf{Y}(P)$ and $\mathbf{N}(P)$ of goals. The sets $\langle \mathbf{Y}(P), \mathbf{N}(P) \rangle$ are duals of the largest *safe* pairs $\langle \mathbf{S}^+(P), \mathbf{S}^-(P) \rangle$ for P, and the latter sets consists of goals that do not *flounder*. Relationships between $\langle \mathbf{R}(P), \mathbf{F}(P) \rangle$, and $\langle \mathbf{Y}(P), \mathbf{N}(P) \rangle$ is summarized as follows. $\langle \mathbf{Y}(P), \mathbf{N}(P) \rangle$ comprise $\langle \mathbf{R}(P), \mathbf{F}(P) \rangle$, and these agree on the sets $\langle \mathbf{S}^+(P), \mathbf{S}^-(P) \rangle$, pairwise. Then completeness of SLDNF-resolution for three-valued validity with respect to sets $\langle \mathbf{S}^+(P), \mathbf{S}^-(P) \rangle$ is shown. The perspicuous proof is based on duality of $\langle \mathbf{S}^+(P), \mathbf{S}^-(P) \rangle$ and $\langle \mathbf{Y}(P), \mathbf{N}(P) \rangle$ and Corollary 4.2.5, Kunen's result for a free term structure, i.e., a non-ground counterpart of a Herbrand structure.

Because of undecidability of non-floundering, researchers proposed several subclasses of non-floundering programs that can be characterized in a syntactic way, e.g. *allowedness*. A wide syntactic class of *μ-correct* logic programs is introduced. This class contains, for example, allowed logic programs. Roughly, a program clause is correct if one can assign modes representing input-output dependencies to arguments $\bar{x}$ in the program clause in a computationally consistent way. Correct goals are shown to be in the set $\mathbf{S}^+$, and hence SLDNF-resolution is complete with respect to correct logic programs and goals.

The final section discusses an asymmetric interpretation of calculi $\mathcal{R}(P)$ and the theories comp(P). Namely each relation R splits into two, R^+ and R^- for positive and negative occurrences, respectively. Calculi $\partial(P)$ are introduced for the asymmetric theory comp$^\sharp(P)$. The asymmetric interpretation is a faithful interpretation from identity-free $\mathcal{R}(P)$ into $\partial(P)$. The partial completion comp$^\sharp(P)$ is a weak theory in the sense that the theory does not prove many properties of logic programs expressed as non-positive formulas. Instead comp$^\sharp(P)$ expresses only that R^+, R^- are closed under (simultaneous) inductive definitions declared in a logic program P. A powerful theory, the *inductive extension* ind$^\sharp(P)$ is defined by saying that R^+, R^- are least fixed points of inductive definitions. Although ind$^\sharp(P)$ proves non-positive formulas that are not provable in comp$^\sharp(P)$, the former is conservative over the latter with respect to positive formulas since positivity implies monotonicity. Some examples are given demonstrating the usefulness of these theories and their correctness is shown.

Errors. Page 648, line 12; page 655, last line; page 657, line 15: These are included in Chapter I of the ***Handbook***. Page 660, formulas (2) and (4): '$\mathcal{R}(P) \vdash$' should denote identity-free derivability.

Toshiyasu Arai

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

Robert L. Constable. *Types in logic, mathematics and programming*. Ibid., pp. 683–786.

Chapter X by R. L. Constable is the second-longest chapter. I regret to say that I am a complete stranger to this area. My comments have to be very brief, but this by no means intends to assert that the subject is less important than ones in other chapters.

In the beginning it is claimed that a new branch of mathematics to be called *formal mathematics* is emerging from the use of automated reasoning and digital libraries. The theorems of this mathematics are completely formal. Completely formal (computer-generated) proofs support the highest known standards of rigor and truth. Creating them require expressive high-level programming languages. Type theory has provided inspirations for designing such languages. According to the author, type theory is simultaneously a *foundational* and *computational* theory. It provides definitions of the basic notions in logic, mathematics, and computer science in terms of a few primitive concepts. Among the primitive built-in concepts in type theory are notions of algorithms, data type, and computation.

Section 2 covers *Typed logic*, and discusses a typed presentation of the predicate calculus in the style of Martin-Löf's type theory as expressed in the system Nuprl. Section 3 covers *Type theory* and is an introduction to Martin-Löf semantics. Section 4 covers *Typed programming languages*. The notion of partial types common in programming is explained. Further, the recursive types and very dependent types are discussed.

Errors. Page 693, line 11 from below: $L(a, x)$ should be $L(x, a)$.

TOSHIYASU ARAI

Faculty of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima 739-8521, Japan. arai@mis.hiroshima-u.ac.jp.

HELMUT SCHWICHTENBERG. *Finite notations for infinite terms*. ***Annals of pure and applied logic***, vol. 94 (1998), pp. 201–222.

One of the aims of proof theory is to use the syntactical information given in a well-formed formula or a well-formed derivation to make analyses about what can and what cannot be done in certain formal systems. There are many ways to give such syntactical information. Some of the most important ones are (1) Gentzen's system of natural deduction and of sequential calculi, (2) Kleene's system O of ordinal notation, (3) Veblen–Bachmann–Ackermann–Schütte–Takeuti systems of ordinal notation, (4) Martin-Löf's theory of types, and (5) the Novikoff–Schütte–Tait analyses of arithmetic using ω-rules. The syntactical objects are either finite or infinite in a uniform way.

Wilfried Buchholz in an important paper (*Notation systems for infinitary derivations*, ***Archive for mathematical logic***, vol. 30 (1991), pp. 277–296) showed how to give notation systems for infinitary sequent-calculus derivations. An essential feature is that from a notation one can find in a primitive recursive way such information as "last rule used." Furthermore the terms are ordered in such a way that the syntactical transformation for getting cut-free derivations produces smaller terms.

The present paper does much the same thing for natural deduction systems (typed lambda-calculi) as Buchholz did for sequent-calculi. This turns out to be more complicated than first expected. The assignment goes in three steps. First we assign infinitary terms as in the Novikoff–Schütte–Tait analyses. Then we code them in a primitive recursive way using the uniformity in the infinitary terms. This is cumbersome, but more or less straightforward. The last step is new. Say we have a notation system $\mathcal{T}$ using primitive recursive codes for the infinitary terms. We extend $\mathcal{T}$ to a new notation system $\mathcal{T}^*$ where we are also able to express directly that we have beta-conversion and reduction among terms. This is done by having new symbols for each variable, and function symbols for beta-conversion and for reduction.

The terms in $\mathcal{T}$ are ordered by the usual tree-ordering of the infinitary terms. In $\mathcal{T}^*$ this is extended to an ordering that also respects reductions among terms. This is done by a straightforward construction but where we get a jump up to the next epsilon-number.

So much for the notation systems; now consider the applications. First consider the system PCF_α of partial recursive functionals with fixed-point operator. In PCF_α we use approximation of functionals. Using the theory, the author is able to prove in a direct and perspicuous way a trade-off theorem, where higher-type recursion is traded off for simpler recursion over longer orderings. The second application is to show continuous normalization for natural deduction. Both of these applications are known. But they are notoriously hard to prove in a natural way. With the new proof here, this difficulty is resolved.

The author shows how the new tools in proof theory given by Buchholz for sequential calculi can now also be used for systems of natural deduction (typed lambda-calculi).

HERMAN RUGE JERVELL

Institute for Linguistics and Philosophy, University of Oslo, P.O. Box 1102, N-0317 Oslo, Norway. herman.jervell@ilf.uio.no.

ANTON SETZER. *Well-ordering proofs for Martin-Löf type theory*. ***Annals of pure and applied logic***, vol. 92 (1998), pp. 113–159.

The paper under review deals with the question of the precise proof-theoretic strength of Martin-Löf's type theory with W-type and one universe ($\mathrm{ML_1W}$ for short). This problem was completely solved in the author's dissertation, and the present article is a rewritten and improved version of the most difficult part of that thesis. It is shown that $\mathrm{ML_1W}$ proves the scheme of transfinite induction up to each ordinal less than $\psi_{\Omega_1}\Omega_{\mathrm{I}+\omega}$. This result can be used to prove that every Π^0_2-theorem of the system $\mathrm{KPi}^{(+)} := \bigcup_{n\in\mathbb{N}} \mathrm{KPi}^+_n$ of set theory is provable in $\mathrm{ML_1W}$ (see e.g. Anton Setzer, *Translating set theoretical proofs into type theoretical programs*, ***Computational logic and proof theory***, Lecture notes in computer science, vol. 1289, Springer, 1997, pp. 278–289). KPi^+_n is Kripke–Platek set theory with one recursive inaccessible (i.e., an admissible closed under the step to the next admissible) and $n-1$ admissibles above it. On the other side, as shown in the author's thesis (or independently in Griffor and Rathjen, *The strength of some of Martin-Löf type theories*, ***Archive for mathematical logic***, vol. 33 (1994), pp. 347–385), $\mathrm{ML_1W}$ can be interpreted in $\mathrm{KPi}^{(+)}$ preserving arithmetic formulas (essentially). Altogether it follows that $\mathrm{KPi}^{(+)}$ and $\mathrm{ML_1W}$ prove the same Π^0_2-sentences, and thus have the same provably recursive functions.

In Section 1 (working in ZFC + existence of a weakly inaccessible I) the author defines a system of collapsing functions $\psi_\kappa : On \to \kappa$ (κ uncountable regular cardinal) and proves its main properties. The system is a slight modification of that presented in Buchholz, *A simplified version of local predicativity*, ***Proof theory***, Cambridge University Press, 1992, pp. 115–147.

In Section 2, a primitive recursive system $(\mathsf{OT}, \prec)$ of ordinal notations is developed on the basis of the ψ_κ's. The elements of OT (called *ordinal terms*) are terms of the forms 0, I, Ω_a, $\psi_\kappa a$, $a+b$, φab with $a, b \in \mathsf{OT}$ and $\kappa \in \{\mathrm{I}\} \cup \{\Omega_{a+1} \mid a \in \mathsf{OT}\}$. $\prec$ is the order induced on OT by the canonical interpretation of the ordinal terms. Important properties of $\prec$ are $\psi_\kappa a \prec \kappa$, $\Omega_b \prec \psi_{\Omega_{b+1}} a$, $a + b \prec \Omega_c \Leftrightarrow \varphi ab \prec \Omega_c \Leftrightarrow \{a, b\} \prec \Omega_c$, $a \prec b \Leftrightarrow \Omega_a \prec \Omega_b$, $\Omega_b \prec \mathrm{I} \Leftrightarrow b \prec \mathrm{I}$.

In Section 3, two versions of $\mathrm{ML_1W}$ are introduced: $\mathrm{ML_{[TD]}}$ and $\mathrm{ML_J}$. The former is essentially the version in the book by Troelstra and van Dalen (***Constructivism in mathematics***, Volume II, North-Holland, 1988); in the latter the elimination rules for the identity type are formulated by using the (additional) constructor J. Since the precise relationship between $\mathrm{ML_J}$ and $\mathrm{ML_{[TD]}}$ is not known, the author carries out his well-ordering proof simultaneously in both systems. Actually he is working in auxiliary systems $\mathrm{ML_{J,aux}}$, $\mathrm{ML_{[TD],aux}}$ where one can more easily switch between the universe and the main level. A precise description of how every statement in $\mathrm{ML_{J,aux}}$ (or $\mathrm{ML_{[TD],aux}}$) translates into a statement in $\mathrm{ML_J}$ (or $\mathrm{ML_{[TD]}}$, respectively) is given in Section 3.3. In Section 3.4 basic set-theoretic notions and abbreviations are introduced in the type-theoretic framework. The most important is the distinction between "classes" and "sets" (of natural numbers). Classes are properties of natural numbers. More formally, one says "$(x)A$ is a class" (and writes $(x)A : \mathrm{Cl(N)}$) for the dependent judgment $x : \mathrm{N} \Rightarrow A$ type. If t is of type N and $(x)A$ is a class, then $t \in (x)A$ abbreviates $A[x := t]$. Sets are objects of type $\mathrm{N} \to \mathrm{U}$ (where U is the universe). For each set t, t^{Cl} denotes its associated class $(y)\mathrm{T}(ty)$ (where $\mathrm{T}(a)$ is the type "named" by a), and $s \in t$ stands for $s \in t^{\mathrm{Cl}}$.

Section 4 contains the well-ordering proof itself. In this section all complicated type-theoretic definitions are suppressed. Instead assumptions are made about possible constructions, which are actually carried out in Section 5. The main technique used in Section 4 is that of distinguished sets ("ausgezeichnete Mengen"), first introduced in Buchholz, *Normalfunktionen und konstruktive Systeme von Ordinalzahlen*, ***ISILC, Proof theory symposion***, Lecture notes in mathematics, vol. 500, Springer, 1975, pp. 4–25. It might be useful to complement the author's explanations and elaborate a little bit on this technique here.

The starting point is the notion of the accessible part of a binary relation R on a class A: $\mathrm{Acc}_R(A) := \bigcap\{X \mid \forall x \in A(\forall y(yRx \rightarrow y \in X) \rightarrow x \in X)\}$. For $A \subseteq \mathsf{OT}$ let $\mathrm{Acc}(A) := \mathrm{Acc}_{\prec\restriction A}(A)$; further we set $\mathrm{Acc} := \mathrm{Acc}(\mathsf{OT})$. As an immediate consequence of its definition, Acc is well founded (with respect to $\prec$). Therefore one can use $\prec$-induction over Acc to prove that Acc is closed under $+$ and φ. Since clearly $0 \in \mathrm{Acc}$, it follows that the subsystem OT_{Γ_0} of all ordinal terms built up from 0 by $+$ and φ is contained in Acc, and hence well founded. In order to prove the well-foundedness of a larger subsystem of $(\mathsf{OT}, \prec)$, especially one that contains terms built with the collapsing function ψ, one has to get a grip on the higher parts of OT, i.e., terms $a \succ \Omega_1$. This can be achieved by the following iteration of the accessible part definition: For $s \in \mathrm{Acc}_0 := \mathrm{Acc} \cap \Omega_1$ let $\mathrm{W}_s := \mathrm{Acc}(\mathrm{C}^{\Omega_s}(\mathrm{W}_{\prec s})) \cap \Omega_{s+1}$ where $\mathrm{W}_{\prec s} := \bigcup_{x \prec s} \mathrm{W}_x$, and (for any $A \subseteq \mathsf{OT}$ and $a \in \mathsf{OT}$) $A \cap a := \{x \in A \mid x \prec a\}$ and $\mathrm{C}^a(A) :=$ closure of $\{0, \mathrm{I}\} \cup (A \cap a)$ under $+$, φ, Ω, and $(\kappa, y) \mapsto \psi_\kappa y$ $(a \prec \kappa)$. Then, for $\mathrm{W} := \bigcup_{s \in \mathrm{Acc}_0} \mathrm{W}_s$, one can show that $\mathrm{W} \cap \Omega_{s+1} = \mathrm{W}_s$ and therefore W is well founded. One can further prove that W is closed under $+$, φ, ψ, and $(x \mapsto \Omega_x)_{x \prec \Omega_1}$. This yields the well-foundedness of the respective subsystem $\mathsf{OT}_{\Omega_{\Omega_1}}$, and especially of the segment $\mathsf{OT}_{\Omega_{\Omega_1}} \cap \Omega_1 = \mathsf{OT} \cap \psi_{\Omega_1}\Omega_{\Omega_1}$. Now one observes that in this well-foundedness proof instead of W one could have taken any non-empty subset $A \subseteq \mathsf{OT}$ satisfying $\forall s(\exists x{\in}A(\Omega_s \preceq x) \rightarrow A \cap \Omega_{s+1} = \mathrm{W}_s(A))$ ($\boldsymbol{Ag}(A)$ for short) and $\forall s \in A \cap \Omega_1(\Omega_s \in A)$ where $\mathrm{W}_s(A) := \mathrm{Acc}(\mathrm{C}^{\Omega_s}(A)) \cap \Omega_{s+1}$. Therefore, a set $A \subseteq \mathsf{OT}$ satisfying $\boldsymbol{Ag}(A)$ is called *distinguished*. The main properties of distinguished sets are (i) every distinguished set A is well founded (with respect to $\prec$), its initial segment $A \cap \Omega_1$ equals Acc_0, and A is closed under $+$, φ, ψ; (ii) any two distinguished sets A, B are comparable insofar as one is an end-extension of the other. Now let $\mathcal{W}$ be the union of all distinguished sets. Then $\mathcal{W}$ is a distinguished class (i.e., $\boldsymbol{Ag}(\mathcal{W})$ holds), and it depends on the set-existence axioms available in the current framework how "far" $\mathcal{W}$ is extending in OT, and, consequently, which segments $\mathsf{OT} \cap a$ can be shown to be well founded. In $\mathrm{ML_1W}$, for example, by using $\mathcal{W}$ one can prove the well-foundedness of $\mathsf{OT} \cap a$ for each a below $\psi_{\Omega_1}(\varepsilon_{\mathrm{I}+1})$, the ordinal of $\Delta^1_2\text{-CA} + \mathrm{BI}$. But with that the power of $\mathrm{ML_1W}$ is not completely exhausted. In contrast to $\Delta^1_2\text{-CA} + \mathrm{BI}$ the system $\mathrm{ML_1W}$ proves $\forall x \in \mathrm{C}^{\Omega_s}(A)(\mathrm{C}^{\Omega_s}(A) \cap x \subseteq \mathrm{W}_s(A) \rightarrow x \in \mathrm{W}_s(A))$ and $\forall x \in \mathrm{C}^{\Omega_s}(A)(\mathrm{C}^{\Omega_s}(A) \cap x \subseteq B \rightarrow x \in B) \rightarrow \mathrm{W}_s(A) \subseteq B$ not only for sets A but for arbitrary *classes* A, B. Therefore in $\mathrm{ML_1W}$ in addition to $\mathcal{W}$ also the classes $\mathcal{W}_1 := \mathrm{W_I}(\mathcal{W})$, $\mathcal{W}_2 := \mathrm{W}_{\mathrm{I}+1}(\mathcal{W}_1)$, etc. can be shown to be distinguished, which yields the well-foundedness of $\mathsf{OT} \cap \psi_{\Omega_1}(\Omega_{\mathrm{I}+n})$ for each meta-natural number n.

Actually the author works with a somewhat different (but equivalent) definition of distinguished set (or class, respectively) which has technical advantages (especially in the type-theoretic setting) but seems to be less perspicuous. This definition reads as follows: $A \subseteq \mathsf{OT}$ is distinguished if $\mathrm{W}(A) := \mathrm{Acc}_{\tau^A}(\mathrm{M}(A))$ is an end-extension of A, where $\mathrm{M}(A) := \{x \in \mathsf{OT} \mid x \in \mathrm{C}^x(A)\}$ and $y\, \tau^A x :\Leftrightarrow y \in \mathrm{C}^x(A) \cap x$. In Section 5 among other things it is shown how $\mathrm{W}(A)$ can be defined in $\mathrm{ML_1W}$ by means of the W-type constructor. By direct application of that constructor one obtains the class W^A_1, which in set-theoretic terms is the least class X such that $\forall x, f[\forall y \tau^A x(fy \in X) \rightarrow (x, f) \in X]$. Then $\mathrm{W}(A)$ is defined as the class of all $x \in \mathsf{OT}$ such that $\exists f[(x, f) \in \mathrm{W}^A_1 \wedge \forall (y, g) \in \mathrm{W}^A_1[(y, g) \preceq_{\mathrm{W}} (x, f) \rightarrow \mathrm{Cor}^A_0(y, g)]]$ where $\mathrm{Cor}^A_0(x, f) :\Leftrightarrow x \in \mathrm{M}(A) \wedge \forall y \tau^A x \exists g[fy = (y, g)]$, and $\preceq_{\mathrm{W}}$ is the subtree relation on W^A_1, i.e., the transitive and reflexive closure of $s \prec^1_{\mathrm{W}} (x, f) :\Leftrightarrow \exists y \tau^A x(s = fy)$.

This paper is an important contribution to type theory and ordinally informative proof theory. Though it is very clear and precise (apart from some irritating typographical errors) and almost self-contained, it is not easy to read, as the subject is thorny and difficult to motivate.

Wilfried Buchholz

Mathematisches Institut, Universität München, Theresienstr. 39, D-80333 München, Germany. buchholz@rz.mathematik.uni-muenchen.de.

Handbook of logic in artificial intelligence and logic programming, Volume 3, ***Nonmonotonic reasoning and uncertain reasoning***, edited by Dov M. Gabbay, C. J. Hogger, and J. A. Robinson, with D. Nute, Handbooks of logic in computer science and artificial intelligence and logic programming, Clarendon Press, Oxford University Press, Oxford, New York, etc., 1994, xix + 529 pp.—therein:

MATTHEW L. GINSBERG. *AI and nonmonotonic reasoning*. Pp. 1–33. (Reprinted with revisions from ***Readings in nonmonotonic reasoning***, edited by Matthew L. Ginsberg, Morgan Kaufmann Publishers, Los Altos, Calif., 1987, pp. 1–23.)

DAVID MAKINSON. *General patterns in nonmonotonic reasoning*. Pp. 35–110.

JOHN F. HORTY. *Some direct theories of nonmonotonic inheritance*. Pp. 111–187.

DAVID POOLE. *Default logic*. Pp. 189–215.

KURT KONOLIGE. *Autoepistemic logic*. Pp. 217–295.

VLADIMIR LIFSCHITZ. *Circumscription*. Pp. 297–352.

DONALD NUTE. *Defeasible logic*. Pp. 353–395.

HENRY E. KYBURG, Jr. *Uncertainty logics*. Pp. 397–438.

D. DUBOIS, J. LANG, and H. PRADE. *Possibilistic logic*. Pp. 439–513.

The emergence, over the last twenty years or so, of so-called "non-monotonic" logics represents one of the most significant developments both in logic and in artificial intelligence. These logics were devised in order to represent *defeasible reasoning*, i.e., that kind of inference in which reasoners draw conclusions tentatively, reserving the right to retract them in the light of further evidence.

Formally, this can be captured by considering a relation $\mid\sim$ of logical consequence holding between sets Γ of sentences (of a given language) and sentences φ (of that language). Statements of the form $\Gamma \mid\sim \varphi$ are taken to express the fact that φ is warranted on the basis of premisses in Γ. The relation $\mid\sim$ is called *monotonic* (in its first argument) if $\Gamma \mid\sim \varphi$ always implies $\Delta \mid\sim \varphi$, whenever Γ is a subset of Δ.

The relation $\models$ of logical consequence of standard, first-order logic (and of many other logics as well) is monotonic in this sense. Thus, it appears that defeasible reasoning cannot quite be formalized in any logic having a monotonic consequence relation: once a conclusion φ is warranted on the basis of a given set Γ of premisses, there is no "retracting" φ, no matter how many additional premisses we put into Γ, i.e., no matter how much "further evidence" we can gather. Hence there is the need to develop alternative frameworks in which such inferences can be adequately formalized.

Of course, it is quite a tall order to supply a relation of logical consequence that is in some sense alternative to the one of classical first-order logic. One has to measure oneself against the stunning success of first-order logic and many of its extensions and variations, beginning with the work of Frege and Dedekind all the way up to the most rarefied reaches of model theory and proof theory.

Monotonicity is, in a sense, the birthmark of first-order logic. Beginning with Frege's ***Begriffsschrift***, first-order logic was developed for the purpose of formalizing mathematical reasoning, and of course there is no retracting conclusions reached in the course of such an enterprise. But since the very beginning there was also another streak in the history of modern symbolic logic, which has to do with the many instances in which symbolic logic is used to formalize everyday reasoning, and indeed to a large extent, successfully so. Therefore, the emergence of non-monotonic logics over the course of the last two decades can also be viewed as the surfacing of a trend that has been there all along in the development of the field.

Pioneering work in the field of non-monotonic logics was carried out by (among others) J. McCarthy, D. McDermott and J. Doyle, and R. Reiter. The realization (which was hardly new) that ordinary first-order logic was inadequate to represent defeasible reasoning was for the first time accompanied by several proposals of formal frameworks within which one could begin to talk about defeasible inferences in a precise way, with the long-term goal of providing for defeasible reasoning an account that could at least approximate the degree of success of modern symbolic logic in the formalization of mathematical reasoning. The publication of a monographic issue of the journal ***Artificial intelligence*** in 1980 (vol. 13 no. 1-2) can be regarded as the "coming of age" of defeasible formalisms.

The excellent volume edited by Gabbay, Hogger, and Robinson can be viewed as an updated version of that seminal 1980 issue of ***Artificial intelligence***, and will be a standard reference for years to come. The book also serves to fill a lacuna in the widely popular and influential ***Handbook of philosophical logic***, also edited by D. Gabbay together with F. Guenthner for the D. Reidel publishing house. In the planning stages of the ***Handbook of philosophical logic***, a decision was made to leave non-monotonic reasoning out of the third volume (subtitled ***Alternatives to classical logic*** and eventually published in 1986; cf. JSL LVII 752–754), probably because the field did not appear to be mature enough for such an inclusion. Obviously, things are quite different now.

The volume edited by Gabbay, Hogger, and Robinson comprises the nine contributions listed at the beginning of this review. The book also contains an informative preface by the main editor, Dov Gabbay, and an extensive index.

The first contribution, by Matthew Ginsberg, serves also as a general introduction to the range of phenomena to be modeled, collectively referred to as defeasible reasoning. One such example is *non-monotonic inheritance*: whenever we have a body of knowledge taxonomically organized, we presuppose that subclasses inherit properties from their superclasses; dogs have lungs because they are mammals, and all mammals have lungs. There can be exceptions, however, which can interact in complex ways. Mammals, by and large, do not fly; bats are an exception, and they do fly; baby bats are exceptional bats, in that they do not fly. (Does that make them unexceptional mammals?) When we infer that Stellaluna, being a baby bat, does not fly, we are resolving all these potential conflicts based on a *specificity* principle: more specific information overrides more generic information.

Another example is the *closed world assumption*: you need to travel from Oshkosh to Minsk, so you consult your travel agent, who, not surprisingly, tells you that there are no direct flights. How does the travel agent know? In a sense, he does not: his database does not list any direct flights between Oshkosh and Minsk, and he assumes that his database is *complete*.

In both these examples, we have an attempt to *minimize* the extension of a given predicate ("exceptional-specimen" in the first case, "flight-between" in the second), and a moment's thought reveals that such a minimization needs to take place not with respect to what the database explicitly contains but with respect to what it implies.

Ginsberg goes on to review the main approaches to the problem of formalizing such inferences, indicating what the main issues are (how to handle conflicting default inferences) and how the so-called "quantitative" or "probabilistic" approaches differ from the "qualitative" or "symbolic" ones. It is worth mentioning here that a major difference in this respect is the fact that qualitative approaches all strive, in one way or another, to achieve some form of *cut*: if φ can be defeasibly inferred from $\Gamma \cup \{\psi\}$, and ψ can be defeasibly inferred from Γ, then φ can also be defeasibly inferred from Γ. The import of cut, as is well known from proof theory, is that the length of an inference does not matter for the cogency of the conclusion, and we are free to adjoin our conclusions to our original premisses without any *increase* in inferential strength. Notably, this principle fails for probabilistic approaches, in which the probability associated with a conclusion decreases with the length of the inference.

The second contribution in the book, by David Makinson, is noteworthy because it focuses on the abstract properties of a defeasible consequence relation. We know that monotonicity will have to go, but what about the other structural properties? Going back to pioneering work by Gabbay, Makinson points out the importance of the *converse of cut*, a principle referred to as "cautious monotony," and that in classical logic it is subsumed under monotony proper: if both φ and ψ can be defeasibly inferred from Γ, then φ can be defeasibly inferred from $\Gamma \cup \{\psi\}$. In other words, we can put our theorems back into our premiss set with no *decrease* of inferential power. Makinson argues that this principle is crucial if we want a well-behaved relation of defeasible consequence, and proceeds to show that it fails in many non-monotonic inference formalisms.

The third contribution, by John Horty, is a survey of a variety of approaches to non-monotonic inheritance networks. Such networks are collections of nodes and directed ("is-a") links representing taxonomic information. When exceptions are allowed, the network is interpreted *defeasibly*. Although the language of non-monotonic networks is expressively limited by design, such networks represent an extremely useful setting in which to test one's intuitions and methods for handling defeasible information. Horty does a particularly good job at explaining the difference between "cautious" and "bold" approaches to defeasible reasoning (these approaches are also sometimes referred to as "skeptical" and "credulous," respectively). This chapter in many ways brings order to a varied field in which it is not always easy to discern the relationships between the different approaches.

David Poole's chapter on Reiter's default logic surveys what is perhaps the most flexible of all the tools devised to represent defeasible inference. A *default* is a *defeasible inference rule* of the form

$$\frac{\alpha : \beta}{\gamma},$$

(for formulas α, β, γ): if α is known, and there is no evidence that β might be false, then the rule allows the inference of γ. The sentences α, β, and γ are respectively known as the prerequisite, justification, and conclusion of the default.

As is clear, application of the rule requires that a consistency condition be satisfied, and rules can interact in complex ways. In particular, it is possible that application of a rule might cause the consistency condition to fail (as when β is $\neg\gamma$). Reiter's default logic uses the notion of an *extension* to make precise the idea that the consistency condition has to be met both before and after the rule is applied. Given a set Γ of defaults, an extension is, roughly (and in typical circular fashion), a maximal subset Δ of Γ whose conclusions both imply all the prerequisites of defaults in Δ and are consistent with all the justifications of defaults in Δ.

Default logic is intimately connected with certain versions of modal logic, known as *autoepistemic logics*, which are the subject of Kurt Konolige's chapter. The idea is that we can sometimes reach conclusions about the state of the world using facts concerning our own knowledge or ignorance. For instance, I can conclude that I do not have a sister given that if I did I would probably know about it, and nothing to that effect is present in my "knowledge base." Of course, such a conclusion is defeasible, because there is always the possibility of learning new facts.

In order to make these intuitions precise, consider a modal language in which the necessity operator L is interpreted as "it is known that." The following abbreviations are useful: given a set S of sentences, $\overline{S}$ denotes the complement of S, LS denotes $\{L\varphi : \varphi \in S\}$, $\neg S$ denotes $\{\neg\varphi : \varphi \in S\}$, and S_0 denotes the subset of S composed of those sentences containing no occurrences of L. The central notion in autoepistemic logic is that of an *extension* of a theory S, i.e., a consistent and self-supporting set of beliefs that can reasonably be entertained on the basis of S. Formally, an extension of S is a theory T satisfying

$$T = \{\varphi : S \cup LT_0 \cup \neg L\overline{T_0} \vdash_{\mathrm{K45}} \varphi\}.$$

In the definition, LT_0 is a set of "introspective axioms" expressing what is known; $\neg L\overline{T_0}$ is a set of "negative introspective axioms," expressing what is not known, and $\vdash_{\mathrm{K45}}$ is derivability in the modal system K45 (interestingly, although extensions are closed under S5 consequence, replacing K45 by S5 in the definition collapses autoepistemic logic into monotonic S5). In general, autoepistemic logic provides a rich and interesting framework whose mathematical properties and connections to other non-monotonic formalisms are well worth exploring. The interested reader is referred to Konolige's chapter for further details.

Vladimir Lifschitz's chapter surveys yet another non-monotonic formalism based on tools and ideas derived from classical logic. *Circumscription*, originally introduced by John McCarthy, is based on the intuition that, all other things being equal, extensions of predicates should be *minimal*. This can be connected to the usual examples from non-monotonic reasoning by considering principles such as "all normal birds fly." Here we are trying to minimize the extension of the abnormality predicate, and assume that x is normal unless we have positive information to the contrary.

Formally, this can be represented using second-order logic, as follows. Given predicates P and Q, let $P \leqq Q$ abbreviate $\forall x(Px \rightarrow Qx)$ and, similarly, let $P < Q$ abbreviate $P \leqq Q \wedge Q \nleqq P$. If $A(P)$ is a formula containing occurrences of a predicate P, then the circumscription of P in A is the second-order sentence $A^*(P)$:

$$A(P) \wedge \neg\exists p[A(p) \wedge p < P].$$

This special case of circumscription would seem not to be sufficient in practice, for in many applications one needs to minimize the extension of P, while allowing the extension of certain other predicates $Q_1, \dots, Q_n$ to vary. In other words, what would seem to be needed is the apparently more general circumscriptive formula

$$A(P; Q_1, \dots, Q_n) \wedge \neg\exists p, q_1, \dots, q_n[A(p; q_1, \dots, q_n) \wedge p < P];$$

but as Lifschitz shows, this case can be reduced to the basic one. From the point of view of applications, circumscription has a major shortcoming, namely the absence of a complete inference procedure. It shares this shortcoming with many other non-monotonic representation formalisms, and it is due to the second-order nature of the circumscriptive formula. All the more crucial, then, is one of Lifschitz's own contributions to the field: the identification of conditions on the formula $A(P)$ that are sufficient to guarantee that the circumscriptive formula $A^*(P)$ has a first-order equivalent. For instance, if $A(P)$ is the formula $\forall x(F(x) \rightarrow P(x))$, then $A^*(P)$ is equivalent to $\forall x(F(x) \leftrightarrow P(x))$.

Donald Nute's chapter is a survey of several different systems of non-monotonic reasoning, from Nute's own defeasible logic and Delgrande's conditional logic, to the Geffner–Pearl theory of probabilistic entailment and Pollock's theory of warrant (perhaps the most philosophically motivated of the lot). It is worth mentioning here that Delgrande's conditional logic builds upon classical work of David Lewis and Robert Stalnaker. A non-truth-functional binary connective $>$ is introduced, where $p > q$ represents the counterfactual implication from p to q. Several plausible sets of axioms of increasing strength are considered, and completeness is established with respect to Stalnaker's "selector function" semantics as well as Lewis's "system of spheres" semantics. The connection to non-monotonic reasoning is provided by, among other things, the fact that $>$ fails to be left-monotonic: the inference from $p > r$ to $p \wedge q > r$ is, in general, not warranted.

The last two chapters of the ***Handbook*** are dedicated to probabilistic approaches to non-monotonic reasoning, which belong to a quite different tradition from the other theoretical frameworks found elsewhere in the book. In particular, Henry Kyburg's contribution on uncertainty logics provides an excellent introduction to the topic, with an extensive treatment of conditional probabilities. (The chapter is unfortunately rife with typographical errors, the most potentially misleading of which is in the definition of Jeffrey conditionalization given in the last displayed formula on p. 423.) So-called possibilistic logics constitute the topic of the chapter by D. Dubois, J. Lang, and H. Prade. This is an approach derived from Zadeh's "fuzzy logic," and the authors do a good job explaining how possibilistic logics are connected to Zadeh's original work.

That covers the contents of the book. Perhaps it is not inappropriate to take this opportunity to assess briefly the present development of the field of non-monotonic logic. There seem to be three major issues that face any attempt to develop a logic that is adequate to represent defeasible reasoning. The first is a condition of material adequacy: we want a system of non-monotonic logic to capture a sufficiently broad range of examples, and to do justice to at least the most entrenched of our intuitions on the subject. The second issue has to do with the extent to which the resulting system is "well behaved" from a logical point of view: how easy it is to define a relation of logical consequence and what abstract properties it has. The third set of issues has to do with the complexity (logical, computational, or otherwise) of the resulting framework.

Where do we stand with respect to these three issues, after twenty or more years of investigations into non-monotonic logic? With respect to the first set of issues, the whole development of non-monotonic logic has been driven by the desire to account for a rich and well-chosen array of examples. Of course, no single framework can aspire to be universal in this respect, but in general non-monotonic logics appear to be better motivated and more materially adequate than a host of other logical frameworks.

Quite a bit of work has been carried out in recent years pointing out the extent to which non-monotonic logics have failed or succeeded in coming up with a well-behaved relation of logical consequence. As David Makinson points out in the book under review, practitioners in the field have encountered mixed success. In particular, one abstract property, cautious monotony, appears at the same time to be crucial and elusive for many of the frameworks found in the literature so far. Part of the problem, perhaps, is that work aimed at overcoming this kind of limitation can sometimes be in tension with the requirement of material adequacy, running the risk of pushing non-monotonic logic closer and closer to its monotonic (and better behaved) counterpart.

Finally, the third set of issues appears to be the most difficult among the ones that have been singled out. Non-monotonic logics appear to be stubbornly intractable with respect to the corresponding problem for classical logic. For instance, in the case of default logic, one obvious source of complexity is the ubiquity of consistency checks, and circumscription owes its intractability to the second-order nature of the circumscriptive axiom. But there are other, often overlooked, sources of complexity that are purely combinatorial. For instance, even when the language is restricted to conjunctions of literals (making consistency checks feasible), the problem of determining whether a given default theory has an extension is still highly intractable (NP-complete, to be precise, as shown by Kautz and Selman), seemingly because it requires checking all possible sequences of firings of defaults. Although some important work has been done trying to make various non-monotonic formalism more tractable, this is perhaps the problem on which progress has been slowest.

G. Aldo Antonelli

Logic and Philosophy of Science, University of California, Irvine, CA 92697, USA. aldo@uci.edu.

Grigoris Antoniou. ***Nonmonotonic reasoning.*** With contributions by Mary-Anne Williams. Artificial intelligence series. The MIT Press, Cambridge, Mass., and London, 1997, xiii + 285 pp.

The book constitutes a comprehensive introduction to the subject, written from the author's perspective. It discusses, with varying degree of detail, the following major areas of non-monotonic reasoning: default logic, autoepistemic logic, several variants of the logic of minimal entailment including predicate circumscription, and applications to the semantics of logic programming. Moreover, it addresses certain abstract aspects of non-monotonic logic and belief revision. The author admits that criteria for selection of material were based largely on his personal preferences. Therefore, after he introduced himself as a supporter of default logic, it is no surprise that this area of non-monotonic logic receives the most attention.

Covering so many different topics in one book must have been a matter of considerable effort, and the author deserves credit for performing this task. His ability to explain certain fundamental concepts, most notably the notion of so-called extension in default logic, is well worth mentioning. Extensive narrative material, motivations, and simple examples seem to target uninitiated readers. The book is very nicely edited and reads well—a decided asset for all those who would consider adopting it as a textbook or a recommended text. It should have a positive impact on the popularity of non-monotonic logic among researchers and students interested in this subject.

The phenomenon of non-monotonic reasoning has two aspects: the syntactic and the semantic one. Accordingly, one can arrive at non-monotonic reasoning patterns driven either by syntactic or semantic motivations.

The syntactic approach originates from a need to formalize methods of drawing conclusions from the absence of information. To that end, in addition to axioms and (monotonic) rules of inference, one may adopt rules of the form

$$\frac{\nvdash \varphi \ \mid \ \cdots}{\psi}$$

(meaning if φ is not provable, and ..., then infer ψ) which are non-monotonic in that they can prove more from fewer premisses. But because in systems that admit non-monotonic rules the notion of provability is often dependent on non-provability, defining a deductive closure for such systems is considerably more involved than the same task in monotonic logic. This is usually accomplished by means of the fixed-point equation

$$\Phi(T, E) = E \tag{1}$$

expressed in terms of a consequence operator Φ applied to two sets of statements: T—a set of premisses, the subject of deductive closure, and E—a set of assumptions for verification of provability or non-provability assertions (e.g. for verification of $\nvdash \varphi$), sometimes referred to as the context. The operator Φ is monotonic with respect to T but not necessarily with respect to E. A set E is called a deductive closure of T iff, in addition to containing classical (e.g. propositional) consequences of T, it is a solution of the fixed-point equation. Although such a set obeys all the (non-monotonic) rules of inference of the system, not every set closed under those rules will necessarily form a solution of the equation. In particular, all elements of the solution set E are *supported* in that they are monotonically derivable from T or follow from E by means of the non-monotonic rules of inference. (Hence, the notion of *supportedness* may be viewed as a generalization of provability.) As a result, the fixed-point style definition of the deductive closure does not allow for inferences that would actually remedy the very lack of information the non-monotonic rules were introduced to deal with in the first place. For example, the intuitively straightforward non-monotonic rule

$$\frac{\nvdash \varphi \mid \nvdash \neg\varphi}{\varphi} \tag{2}$$

(meaning if φ is independent from what has been proved so far then infer φ) cannot produce from $T = \emptyset$ any solutions E of the fixed-point equation (1).

For all its natural appeal, such a definition introduces two major complications. First, the fixed-point equation, as opposed to the classical definition of deductive closure via the concept of proof, may provide no clue how to compute its solution. Secondly, the equation may have more than one solution (which, as a matter of fact, increases the expressive power of the logic in question) or no solutions at all (which creates obvious problems of an ontological nature). Moreover, being an implicit definition, it may make the explicit meaning and the deductive properties of the system rather difficult to determine.

Default logic, in addition to standard propositional axioms and modus ponens, allows for non-monotonic rules of inference called *defaults*. They have the form of

$$\frac{\varphi : \Diamond\psi_1, \dots , \Diamond\psi_n}{\chi} \tag{3}$$

(where φ and $\Diamond\psi_i$'s are skipped if tautologically true). The meaning of rule (3) is formally described by

$$\frac{\varphi \in T \mid \neg\psi_1 \notin E \mid \cdots \mid \neg\psi_n \notin E}{\chi \in T}. \tag{4}$$

Given a set D of defaults, the operator $\Phi(T, E)$ (the author uses the notation $\Lambda_T(E)$) of closure of T under rules (3) of D relative to E is defined as the closure of T both under propositional consequence and under the corresponding productions (4) of D. A set E is called an *extension* of T iff it is a solution of (1). As a result, E is both supported and closed under all rules (3) of D.

Autoepistemic logic, instead of non-monotonic defaults, uses axioms expressed in a language $L_\Box$ of modal logic S5. In addition to axioms and rules of S5, it admits the non-monotonic rule of inference

$$\frac{\nvdash \varphi}{\neg\Box\varphi} \tag{5}$$

whose meaning is formally described by

$$\frac{\varphi \notin E}{\neg\Box\varphi \in T}. \tag{6}$$

The necessitation rule is interpreted similarly, that is by

$$\frac{\varphi \in E}{\Box\varphi \in T}.$$

The operator $\Phi(T, E)$ (the author uses the notation $\Omega_T(E)$) of closure of T under the rule (5) and necessitation relative to E is defined as the closure of T under both rules and propositional consequence. A set E is called an *expansion* of T iff it is a solution of (1). As a result, E is supported and satisfies the so-called *stability* conditions:

$$\varphi \in E \quad \text{iff} \quad \Box\varphi \in E \quad \text{iff} \quad \neg\Box\varphi \notin E. \tag{7}$$

The semantic approach to non-monotonic logic restricts the semantics of its language L to selected models for L by means of certain context-sensitive assumptions, such as minimality assumptions. In the latter case, a partial order $<$ is superimposed on the class of semantical structures for L. It yields a definition of $<$-minimal entailment $\vdash_<$ where $T \vdash_< \varphi$ iff φ is true in every $<$-minimal model of T. Examples of $<$ in the class of first-order structures include those defined by proper inclusions $\subsetneq$ between models' respective relations, and by proper inclusion $\subsetneq$ between models' universes, as well as their variants and combinations.

Predicate circumscription aims at defining the concept of relation-minimal model of first-order theory within its language, which goal obviously requires second-order logic. (Similarly domain circumscription aims at defining the concept of Herbrand model, which requires infinitary logic.) This has been accomplished in terms of the so-called second-order circumscription. Its first-order fragment is determined by corresponding minimal entailments $\vdash_<$. Also, some relationships between the logic of minimal entailment and default logic are known. For instance, a minimal model of a set T that admits elimination of quantifiers can be defined in terms of an extension of T and the set of defaults

$$D_{\text{nas}} = \left\{ \frac{: \Diamond\neg\varphi}{\neg\varphi} \mid \varphi \text{ is an atomic sentence} \right\}.$$

On the other hand, minimal entailment seems to evade the schemes defined by the defaults because it is characterized by the inference rule of the form "if adding φ to the premiss set T does not enlarge the set of positive consequences of T then infer φ," which clearly does not fall under the scheme (4) (obviously, not under D_{nas}).

Exposition of default logic is by all means the strongest aspect of the book. The author follows very closely R. Reiter's original paper (*A logic for default reasoning*, ***Artificial intelligence***, vol. 13 (1980), pp. 81–132) and his notation, albeit in slightly different order. In addition to Reiter's theory of defaults, the author covers in detail the following variants of default logic: semi-normal default theory (with defaults of the form

$$\frac{\varphi : \Diamond(\psi \wedge \chi)}{\chi}$$

only), justified and constrained default logic (where the modal operators $\Diamond$ that precede ψ_i's in default rules (3) of set D are *de facto* pulled in front of D), and prioritized default logic (where certain sequences of applications of default rules are not allowed, thus eliminating some extensions from consideration).

Although this part of the book seems to read well, there are certain details that leave room for improvement. Here are examples.

The author skips occurrences of the possibility modal operator $\Diamond$ in front of ψ_i's in (3). This seems to be a source of confusion. In particular, the informal interpretation of (3) on page 26 is incorrect because it actually refers to $\Diamond(\psi_1 \wedge \cdots \wedge \psi_n)$ rather than to $\Diamond\psi_1 \wedge \cdots \wedge \Diamond\psi_n$. The discussion of the meaning of (3) and the formal definition of its applicability on page 27 are flawed as well because they tacitly replace $\varphi \in T$ in (4) with $\varphi \in E$, so that the implicit requirement of the supportedness of extension is lost.

In Chapter 7, *Alternative approaches*, the author continues his tendency to interpret premisses ψ_i in (3) as if they were not in disjoint scopes of the modal operator $\Diamond$ which leads to questionable conclusions. In particular, the author's "joint consistency" postulate reduces (3) to its special form

$$\frac{\varphi : \Diamond(\psi_1 \wedge \cdots \wedge \psi_n)}{\chi}.$$

It remains unclear why "jointly consistent" defaults would be necessary; the author does not show that they cannot be easily expressed by appropriate default rules of the form (3). The motivating example (p. 73) is incomplete; the default theory T considered there does not include any axiom or rule that would allow for inferring $\forall x(\mathit{usable}(x) \supset \neg \mathit{broken}(x))$. Consequently there is no paradox in concluding $\mathit{usable}(a) \wedge \mathit{usable}(b) \wedge (\mathit{broken}(a) \vee \mathit{broken}(b))$, which the author intended as justification for introducing the joint consistency postulate.

The autoepistemic logic part (Chapters 9–11) yields to the default logic part both in clarity and correctness. First, the introductory example (*Are the Rolling Stones giving a concert in Newcastle next week? No, because otherwise I would have heard about it.*) and its discussion are misleading in that they suggest that within autoepistemic logic one can express facts that are valid but unknown. This clearly is not the case by virtue of the stability conditions (7). Consequently, the author's interpretation of the provability operator $\Box$ (the author uses symbol L instead of $\Box$, which produces such ambiguous combinations of symbols as *Laussie*—meaning $\Box$*aussie*) as the "knowledge or belief" operator, as well as his referring to E as the "world view" may further confuse the reader.

The semantical aspects of autoepistemic logic are not adequately explicated in the book. What the author calls in Section 9.2 a "semantics" of autoepistemic logic is a halfway measure since the meaning of $\Box$ is defined there in purely syntactic terms as $(v, E) \models \Box\varphi$ iff $\varphi \in E$, where v is a propositional valuation, as opposed to, for example, Kripke semantics of S5. This is an unfortunate omission because every consistent expansion—being a stable set—is equal to the theory of a certain Kripke model for S5. (Kripke-style semantics for autoepistemic logic has been investigated elsewhere; cf. Kurt Konolige, *On the relation between default and autoepistemic logic*, ***Artificial intelligence***, vol. 35 (1988), pp. 343–382, and Wiktor Marek and Mirosław Truszczyński, *Autoepistemic logic*, ***Journal of the Association for Computing Machinery***, vol. 38 (1991), pp. 588–619.)

Chapter 11, *Embedding default logic into variants of AE-logic*, explores the idea of expressing defaults (3) by means autoepistemic formulas

$$(\Box\varphi \wedge \Diamond\psi_1 \wedge \cdots \wedge \Diamond\psi_n) \supset \chi. \tag{8}$$

This task is far from being trivial since the meaning (4) of rule (3) is clearly not the same as the meaning of formulas (8); for example, the meaning of the premiss of (8) does not depend on T. The author begins with three unsuccessful attempts (Sections 11.1–11.3) which are supposed to condition the reader for the fourth one (Section 11.4) based on the concept of *strongly grounded expansion*, only to end up with a false result (Theorem 11.4) which claims that strongly grounded expansions translate exactly on extensions. A counterexample to this claim was published by Wiktor Marek and Mirosław Truszczyński (*Relating autoepistemic and default logics*, ***Proceedings of the first international conference on principals of knowledge representation and reasoning***, Morgan Kaufmann, 1989, pp. 276–288). A serious student of autoepistemic logic may find Marek and Truszczyński, *Autoepistemic logic* (cited earlier), an easier (and far more complete) text to follow.

Chapter 12, *Circumscription*, focuses almost entirely on old attempts to express second-order properties with first-order formulas, and ignores certain more successful approaches to the minimization problem.

In the spirit of early misconceptions about the expressive power of circumscription, the author begins with an introductory example whose goal is to deduce from a theory

$$T = \{\mathit{bird}(\mathit{Tweety}),\ \forall X(\mathit{bird}(X) \wedge \neg\mathit{abnormal}(X) \supset \mathit{flies}(X))\}$$

a positive formula $\mathit{flies}(\mathit{Tweety})$ which the vanilla-style circumscription cannot derive (a variable circumscription is necessary to achieve this, as on p. 154). Later on, the author attempts to accomplish that effect without use of circumscription by "minimizing" the extension of predicate *abnormal* by means of additional axiom $\forall X \neg\mathit{abnormal}(X)$. This, most likely, will mislead many readers.

Despite the fact that circumscription and various versions of minimal entailment appear to be the technically most advanced and studied form of non-monotonic reasoning, the author consistently downplays their importance. He denies circumscription any substantial application or deeper theoretical results and mischaracterizes its expressive power. All this is in spite of the fact that he uses the concept of a minimal Herbrand model (whose only definition has been hidden within Theorem 18.1 some hundred pages after Section 12.2, *Minimal models*) to define the stable semantics of a logic program.

The author's remarks about the allegedly intrinsic nature of incompleteness of circumscription (pages 151–152) are largely irrelevant in that they apply only to its earliest version. His suggestion stating that "Usually other variants of circumscription found in the literature are incomplete if a first-order schema is used" in effect dismisses numerous counterexamples (cf. Marek Suchenek, *Preservation properties in deductive databases*, ***Methods of logic in computer science***, vol. 1 (1994), pp. 315–338).

The author turns to a critique of Reiter's version of the closed-world assumption, ignoring later development in this subject by Jack Minker (*On indefinite databases and the closed world assumption*, ***6th conference on automated deduction***, Lecture notes in computer science, vol. 138, Springer-Verlag, 1982, pp. 292–308) and others, which resolved all early paradoxical properties of its generalizations. A section devoted to prioritized circumscription offers no results, the most basic of which would be the ability to infer positive conclusions that vanilla-style circumscription could not derive. Although prioritized circumscription provides some of the standard semantics for stratified logic programs, no connections between the two are mentioned.

Part IV, devoted to abstract non-monotonic logic, lacks coherence with the rest of the book. For example, in Section 13.2, page 166, there is a list of postulates describing what the author refers to as "natural and useful ways of reasoning." One of them is *cumulativity*, which the flagship logic of the book (default logic, that is) does not satisfy (as noted on p. 74), except in a very special case of "supernormal" theories. The same postulate is not satisfied by stable model semantics, discussed in detail later in the book.

Section 13.5, *Preferential models*, does not relate to default logic or to autoepistemic logic because the author does not discuss their semantics. Moreover, preferential model theory does not apply to circumscription (or to the logic of minimal entailment $\vdash_<$) either, because of violation of the *rational monotony* postulate (if $\{\varphi\} \not\mid\!\sim \neg\psi$ and $\{\varphi\} \mid\!\sim \chi$ then $\{\varphi \wedge \psi\} \mid\!\sim \chi$). (Here is a counterexample. Let $\varphi = p \vee q$, $\psi = p \vee r$, and $\chi = \neg r$. We have $\{\varphi\} \nvdash_< \neg\psi$ and $\{\varphi\} \vdash_< \chi$, but $\{\varphi \wedge \psi\} \nvdash_< \chi$.) Therefore, one cannot be surprised that the minimal entailment, which is widely accepted as the classic instance of preferential semantics, is not mentioned in that part.

Chapter 14, *Belief revision*, incorporates overly complicated notation for rather elementary notions. For example (pp. 186–187): "formally, an expansion function $^+$ is a function from $\mathcal{K}_\mathcal{L} \times \mathcal{L}$ to $\mathcal{K}_\mathcal{L}$, mapping (T, φ) to T^+_φ where $T^+_\varphi = \mathrm{Th}(T \cup \{\varphi\})$"—quite an elaborate definition of the set of consequences of $T \cup \{\varphi\}$. A motivating example (p. 189) attributed to Hansson actually contradicts axiom $(^-8)$ of "a contraction function $^-$" whose existence is not even proved in the book. Page 192 uses the notation "$\varphi \wedge \psi = \varphi$," leaving the reader to discover that what is meant there is that the formulas are "equally entrenched."

In Chapter 15, *Implementing belief revision*, the central concept of *finite partial entrenchment* is not defined. The author uses instead a *finite partial entrenchment ranking* **B**, for a measurement of the level of "epistemic importance" of a statement, the adequacy of which is not expounded. For instance, it follows from that definition that if $\mathbf{B}(\varphi)$ (the level of belief in φ) is > 0 then $\mathbf{B}(\neg\varphi) = 0$, provided that $\neg\varphi$ and *false* are in dom($\mathbf{B}$). Despite this fact, the author attempts to explain the meaning of $\mathbf{B}(\varphi)$ in terms of probability.

Part V of the book is devoted to non-monotonic aspects of logic programming. At the outset of Chapter 17, *The system Theorist*, in which the author describes Poole's system (implemented in Prolog) for non-monotonic reasoning, one can find the following characterization of relationships between various non-monotonic logics: "Default and Autoepistemic Logic go beyond classical predicate logic, whereas Circumscription in its simple forms works within predicate logic, but changes the way reasoning is treated by adding new premises" (p. 229). This enunciation will, most likely, mislead many readers; default logic and autoepistemic logic only "go beyond" propositional calculus (sentences of the form $\forall x \Box\varphi(x)$ or $\forall x \Diamond\varphi(x)$ are not allowed in the language of autoepistemic and default logic), and "adding new premises" alone will not make circumscription non-monotonic. Then he reservedly cites Poole who "argued that there is nothing wrong with first-order logic."

The idea behind *Theorist* is somewhat similar to that of model-theoretic forcing. There is a first-order theory T and a set Δ which comprises conditions p for T (sets of quantifier-free sentences), with a requirement that $T \cup p$ be consistent. Maximal conditions G play the role of generic sets for T. They define extensions, that is, the sets of first-order consequences of $T \cup G$. Intersections of extensions may be thought of as sets of (forcible) conclusions. A version of the above construction with the relation of forcing defined by usual induction (e.g. $p \Vdash \neg\varphi$ iff for no $q \supseteq p$ in Δ, $q \Vdash \varphi$) and generic sets as the ones that force every sentence or its negation, has been known to coincide with the minimal entailment for sets of clauses. Since forcing can be characterized in terms of not enlarging the set of certain consequences of T (a property not expressible, except for special cases, by default rules), the author's unproved claim that the logic of *Theorist* is reducible to default logic seems doubtful. (What the author actually demonstrates—Theorem 17.1—is much less: every extension in the sense of *Theorist* is an extension in the sense of default logic with certain defaults of the from

$$\frac{:\Diamond\varphi}{\varphi},$$

and vice versa.)

Chapters 18 and 19 are devoted to semantics of logic programs, and constitute rather *verbatim* renderings (as the author acknowledges, in the case of Chapter 19) of a number of classic papers in this subject. Providing some perspective would greatly benefit the readers. For instance there is no mention that "classical negation" in program clauses may be reduced to a non-monotonic one by a rather simple technical trick. Stratified programs as well as disjunctive ones, and their relation to prioritized circumscription are not adequately explained. A classic result from Geneviéve Bossu and Pierre Siegel (*Saturation, non-monotonic reasoning, and the closed-world assumption*, ***Artificial intelligence***, vol. 25 (1985), pp. 13–63), is presented in a weaker form (Theorem 18.1) without attribution.

Notwithstanding its weaknesses, publication of this work benefits those willing to get acquainted with non-monotonic logic, provided they will arm themselves with a dose of skepticism and follow the author's own suggestion not to take all his claims and opinions for granted.

MAREK A. SUCHENEK

Department of Computer Science, California State University at Dominguez Hills, Carson, CA 90747, USA. suchenek@csudh.edu.

MAARTEN MARX and YDE VENEMA. ***Multi-dimensional modal logic.*** Applied logic series, vol. 4. Kluwer Academic Publishers, Dordrecht, Boston, and London, 1997, xiii + 239 pp.

Multi-dimensional modal logic—MDML—is a new branch of modal logic and, as stated in the *Editorial preface*, the book "is the first comprehensive coverage of the frontline research of a topic which is now central to applied logic. Multi-dimensional modal logics have applications in such diverse areas as the logical analysis of language, deductive and temporal databases, fibring and combining systems, and the semantics of generalised quantifiers."

One of the main applications of the book, however, is connected to MDML's close relationship to algebraic logic in the sense of Tarski. Tarskian algebraic logic studies algebraic systems related to the theory of quantification and to the theory of binary and n-ary relations, and thus connects with many-dimensionality. This field has a long history and is rich in deep theories and results. One aspect of MDML can be viewed as an application of the methods and intuitions from modal logic to the old area of algebraic logic. Conversely, MDML takes ideas and methods from algebraic logic, adapting them into a suitable modal form. In my view, a key feature of the book is the systematic demonstration of the fruitfulness of the "modal approach" to some problems of algebraic logic and the "algebraic logic approach" to some problems of MDML.

The aim of the authors is not only to present a series of multi-dimensional modal formalisms, but also to present and illustrate a number of specific methods typical in the field, such as the "step-by-step method" and "the mosaic method" of constructing "standard" semantical structures, axiomatizing systems by unorthodox "Gabbay-style rules," and representation theorems of abstract semantical structures as concrete multi-dimensional structures. The book itself is a good fusion of the authors' dissertations, each on some special subfield of the area: *Many-dimensional modal logic* by Y. Venema, Amsterdam 1992, and *Algebraic relativization and arrow logic* by M. Marx, Amsterdam 1995.

What is multi-dimensional modal logic? Following Kripke, the main semantical structures for modal logics are relational systems called frames, consisting of a non-empty set, whose elements are called "possible worlds" or "states," and one or more relations on this set, called "accessibility relations," used for the interpretation of the modal operators. In such structures the atomic formulas are interpreted in each state by "true" or "false" and then this is extended inductively for arbitrary formulas. For instance, if there is only one modality $\Diamond A$—the possibility operator (called "diamond")—the corresponding semantical structure is of the form (W, R) where R is a binary relation on W. Then the interpretation of $\Diamond A$ is the following: for $x \in W$,

$$x \Vdash \Diamond A \quad \text{iff} \quad \exists y \in W \text{ such that } xRy \text{ and } y \Vdash A.$$

Informally: $\Diamond A$ is true at the world x iff A is true at some world y accessible from x by the relation R.

Normally, possible worlds are abstract objects without any internal structure and the accessibility relations are also some abstract relations, satisfying some abstract conditions like reflexivity, transitivity, and so on. For applications, however, one often needs to use concretely defined semantical structures in which the possible worlds have some concrete internal structure and the accessibility relations between the possible worlds depend on that structure. For MDML, possible worlds have a structure of tuples or sequences $s = (s_0, s_1, \dots)$ of some arbitrary but fixed length α, over some base set, say X. Here α is called a dimension, for instance if $\alpha = 2$, MDML is called "two-dimensional modal logic." Given the base set X and a dimension α it is possible to take all α-dimensional sequences or some of them. In the first case the authors use the terminology "full square" semantics for $\alpha = 2$, or "full cube" semantics for $\alpha > 2$. In the second case, the corresponding semantics is called relativized square-semantics or relativized cube-semantics.

There are various ways of defining accessibility relations depending on dimensionality. For instance, in a two-dimensional case one possible binary relation $\equiv_0$ between two 2-tuples (u_0, u_1) and (v_0, v_1) is the following: $(u_0, u_1) \equiv_0 (v_0, v_1)$ iff $u_1 = v_1$. If we define a modal operator $\diamondminus$ having $\equiv_0$ as its accessibility relation, then its semantics will be the following:

$$(u_0, u_1) \Vdash \diamondminus A \quad \text{iff} \quad \exists (v_0, v_1) \text{ such that } (u_0, u_1) \equiv_0 (v_0, v_1) \text{ and } (v_0, v_1) \Vdash A.$$

Or simply: $(u_0, u_1) \Vdash \diamondminus A$ iff $\exists v_0\, (v_0, u_1) \Vdash A$. The intuitive two-dimensional meaning of $\diamondminus$ is that $\diamondminus A$ is true at some point (u_0, u_1) iff A is true somewhere on the "horizontal line" through (u_0, u_1).

It is clear from this example that MDML is a branch of modal logic, studying modal systems with multi-dimensional semantics like those described above. In the abstract possible world semantics, sentences are identified with some subsets of the universe of the model. Here, in the case of MDML, sentences are subsets of some Cartesian product of the universe and hence are interpreted as relations and this makes the connection of MDML with some theories of relations studied in algebraic logic quite natural.

The book contains six chapters, plus two appendices containing all needed definitions and general facts from modal logic. The role of the appendices is, on the one hand, to unify the notations, and, on the other, to be used as a suitable source of references. It is presupposed that the reader is familiar with some basic facts of modal logic.

Chapter 1 is introductory and contains some motivations for the field as a whole, examples from the literature, and a short overview.

One of the most important chapters from a methodological and pedagogical point of view is Chapter 2, *Two-dimensional modal logics*. The main purpose of this chapter is to introduce the reader to the methods and techniques typical for the MDML, by means of some carefully selected and not too complicated examples of modal systems. The two-dimensional case is very suitable for that purpose. The exposition here is more detailed and has many intuitive explanations. I will also be more detailed in representing the content of this chapter.

Section 2.1, *Operations on the square universe*, starts with a list of basic accessibility relations, typical of the two-dimensional case. For a frame (W, I) where $W = U \times U$, I is a set of binary relations, corresponding to some two-dimensional modalities:

$$I(\iota\delta) = \{(u, v) \mid u = v\},$$
$$I(\diamondminus) = \{((u, v), (x, y)) \mid v = y\}, \quad I(\diamondbar) = \{((u, v), (x, y)) \mid u = x\},$$
$$I(\diamondplus) = \{((u, v), (x, y)) \mid u = y\}, \quad I(\diamondplus') = \{((u, v), (x, y)) \mid v = x\},$$
$$I(\otimes) = \{((u, v), (x, y)) \mid u = y, v = x\},$$
$$I(\obar) = \{((u, v), (x, y)) \mid u = x = y\}, \quad I(\ominus) = \{((u, v), (x, y)) \mid v = x = y\},$$
$$I(\circ) = \{((u, v), (w, x), (y, z)) \mid u = w, v = z, x = y\},$$

The semantics of these modalities are visualized by two-dimensional coordinate pictures. For instance $\iota\delta$ is a propositional constant true exactly "on the diagonal." The meaning of $\diamondminus$ has already been explained and $\diamondbar A$ is true at some point iff A is true somewhere on the "vertical" through this point; $\otimes A$ is true at some point iff A is true at the point symmetrical to the "diagonal," and so on. The ternary relation of the last modality is sometimes called "the triangle relation." The corresponding modality is a two-place one, called composition, because its semantics is a modal analogue of the composition of two relations: $(u, v) \Vdash A \circ B$ iff $\exists w, x, y, z \in U$ such that $((u, v), (w, x), (y, z)) \in I(\circ)$ and $(w, x) \Vdash A$ and $(y, z) \Vdash B$; or more simply, which directly shows the analogy with composition:

$$(u, v) \Vdash A \circ B \text{ iff for some } x \in U \text{ both } (u, x) \Vdash A \text{ and } (x, v) \Vdash B.$$

Section 2.2, *Axiomatizing S5-square*, deals with a two-dimensional modal logic of the modalities $\diamondminus$ and $\diamondbar$ interpreted in a full square. The name "S5-square" comes from the traditional modal logic, because the modalities $\diamondminus$ and $\diamondbar$ are of S5 type.

Section 2.3, *Cylindric modal logic of squares*, MLR_2, extends the logic of S5-squares with the identity constant $\iota\delta$ and $\otimes$ interpreted in the full square semantics and is connected with two-dimensional cylindric algebras of binary relations. It is proved that MLR_2 corresponds to the two-variable fragment of first-order logic with identity. Since by a theorem of Scott this fragment is decidable, MLR_2 is also decidable. Completeness theorems for the logic of "S5-square" and MLR_2 demonstrate two simple applications of the step-by-step method.

Section 2.4, *The modal logic of composition*, shows the striking difference between composition and the other two-dimensional modalities. First, by a result of Andréka the modality of composition does not have a finite structural derivation system, complete in the full-square semantics. Secondly, by a remarkable result of Andréka, Kurucz, Németi, Sain, and Simon, the logic of composition on the full square is undecidable.

Section 2.5, *A two-dimensional temporal logic*, axiomatizes a logic complete in the Cartesian product of two linear orderings. The logic has four "compass" modalities $\diamondN$, $\diamondS$, $\diamondE$ and $\diamondW$, which are mnemonic for *north, south, east*, and *west*. Here is the first example in the book of an axiomatization by means of a non-structural derivation system with Gabbay-style irreflexivity rules.

Chapter 3, *Arrow logic:* The basic language of arrow logic (AL) is an extension of the logic of composition and consists of the operators $\{\circ, \otimes, \iota\delta\}$ of composition, converse, and identity. (See the review JSL LXIII 333.) The standard semantics is the two-dimensional one studied in Chapter 2, but here a more general semantics is also proposed, consisting of frames of the form (V, C, F, I), called arrow frames, where V is a non-empty set of abstract objects, called arrows, C is a ternary relation in V for interpreting $\circ$, F is a binary relation in V for interpreting $\otimes$, and I is a subset of V for the interpretation of $\iota\delta$. The structure (V, C, F, I) is called a pair frame if V is a set of pairs of some base set U and $\{\circ, \otimes, \iota\delta\}$ have their standard pair semantics. Where $V = U \times U$, the class of such frames is denoted by SQ (full squares) and where V is a reflexive and symmetric relation in U, the corresponding class is denoted by LSQ (local squares).

The introductory part of the chapter gives some motivations as to why the above operations are the basic operations of arrow logic. It is mentioned that arrow intuition and arrow logic are connected with various abstract and applied areas such as mathematics—vectors, functions, relations, morphisms in category, and so on; computer science—programs; linguistics—dynamic semantics; cognitive science—cognitive actions. A special subsection is devoted to dynamic semantics, but the explanations why just this signature is the adequate one are quite debatable. For instance it is difficult to find analogues of the identity constant $\imath\delta$ and the converse $\otimes$, while the motivation with the theory of relation algebras and the connection with the first-order logic of binary predicates is established and explained very well.

As it was for the case of the logic of composition, AL in its full square semantics (SQ) is undecidable and does not possess a finite structural derivation system, while AL in its local-square semantics (LSQ) is decidable and possesses a finite structural derivation system.

Connections with the theory of relation algebras from algebraic logic are discussed in detail. A special section is devoted to a temporal extension of arrow logic. The chapter ends with some discussion of alternative directions in arrow logic: dynamic arrow logic of van Benthem, Bulgarian-style arrow logic, connections with modal transition logics like PDL—propositional dynamic logic, connections with Lambek calculus, and arrow logics of higher dimensions.

Chapter 4, *Modal logics of intervals:* This chapter contains an application of two-dimensional modal logic to the interval temporal logic. The main idea is very simple: represent an interval in a linear order $(T, <)$ as an ordered pair (s, t) with $s < t$. Then all intervals in T will be two-dimensional points in the "northwestern half-plane" of $T \times T$. In this way the interval tense logic is represented as a two-dimensional tense logic. All meaningful binary relations between two intervals as *is later than*, *meets*, *overlaps*, *ends*, *during*, *begins*, and their converses can be reinterpreted geometrically as possible configurations between two points in the half-plane. As an example, a detailed study of the system HS of Halpern and Shoham is given.

Chapter 5, *Modal logics of relations*, is devoted to an application of the methods of multi-dimensional modal logic to the study of n-ary relations for $n > 2$. This is closely related to the theory of n-dimensional cylindric algebras over full cubes, relativized cylindric algebras, finite-variable fragments of first-order logic, and the theory of quantification. In this approach existential quantifiers are treated as diamonds, the identity of ith and jth coordinates of an n-tuple are special propositional constants $\imath\delta_{ij}$, and substitutions and permutations of the variables are treated by special modalities. Unlike the two-dimensional case, the n-dimensional modal logics of full cubes are not decidable and do not possess finite structural axiomatization. But if we replace the full cube semantics with relativized ones, the corresponding logics become decidable and finitely axiomatizable by orthodox structural derivation systems. Two main examples are considered here—the n-dimensional cylindric modal logic CML_n interpreted over full n-cubes and the logic MLR_n interpreted over (relativized) local cubes. The language of MLR_n extends the language of CML_n with operators for substitutions. Both systems have non-trivial completeness theorems—CML_n by a Gabbay-style derivation system and MLR_n by an application of the method of mosaics.

Chapter 6, *Multi-dimensional semantics for every modal language*, is the last chapter. It contains a generalization of the two-place operation of composition $A \circ B$ of binary relations to an n-place composition of $\bullet(A_1, \dots, A_n)$, of n n-ary relations. The main theorem here is a completeness theorem for an arbitrary n-place normal modality with respect to an n-dimensional semantics. Note, however, that the result is true only for the minimal modal logic of the language.

Now for some remarks. Although the number of the papers on MDML is not large, the bibliography in the book is incomplete. Some additional papers are the following. A. Arsov, *Basic arrow logic with relation algebraic operators*, ILLC research reports and technical notes series, LP-1994-02, 1994, University of Amsterdam; D. Gabbay and V. Shehtman, *Products of modal logics, part 1*, ***Logic journal of the IGPL***, vol. 6 (1998), pp. 73–146; S. Passy and T. Tinchev, *An essay in combinatory dynamic logic*, ***Information and computation***, vol. 93

(1991), pp. 263–332; A. Petkov, *Propositional dynamic logic in two- and more dimensions*, ***Mathematical logic and its applications***, edited by D. Skordev, Plenum Press, 1987, pp. 323–329; V. Shehtman, *Modal logics of domains on the real plane*, ***Studia logica***, vol. 42 (1983), pp. 63–80; D. Vakarelov, *Many-dimensional arrow logics*, ***Journal of applied non-classical logics***, vol. 6 (1996), pp. 303–345.

1. *On the multi-dimensional intuition in the book.* Multi-dimensional objects considered in the book are mainly pairs or n-tuples with fixed n. There are two main geometrical ways to represent pairs—as two-dimensional points in a Cartesian coordinate square and as arrows in a directed graph. The first one is the main geometrical intuition adopted in the book, which is typical for the Amsterdam–Budapest schools and the second is mentioned in connection to the Bulgarian direction on arrow logic. The second representation has an advantage with some generalizations of directed graphs—directed multi-graphs (admitting more arrows between two points) and n-dimensional multi-graphs, called also n-dimensional arrow structures. In the book this graph intuition is missing. The pair intuition in the book is so strong that it even appears where the authors try to represent the graph intuition for the Bulgarian direction of the field, considering intuitively an arrow x as leading from its left point $l(x)$ to its right point $r(x)$, overlooking the fact that "left" and "right" are meaningless for arrows. This misreading obviously comes from the pair intuition—writing x as $x = (x_1, x_2)$ puts x_1 on the left and x_2 on the right.

Let me mention that the pair intuition influences the structure of some proofs. In order to understand the proofs, we have to draw pictures. It is still possible to draw pictures in two-dimensional Cartesian coordinate squares (which is good for full squares and not so good for the relativized ones), while this is almost impossible for higher dimensions. There are no difficulties, however, in drawing pictures for n-dimensional arrows for an arbitrary n. Of course this does not mean that the pair intuition should be neglected—both need to be used according to their appropriateness.

Let me note also that sets of pairs or n-tuples are not the only possible semantics for MDML—other candidates are n-dimensional arrow structures mentioned previously. So MDML can also be considered as a natural logic for arrows, as it is treated in the Bulgarian direction in the field.

2. *On the place of arrow logic in multi-dimension modal logic.* Arrow logic (AL) in this book is a very special case of two-dimensional modal logic, containing in its language the operations of composition and converse and the identity constant, taken from the algebra of binary relations. The motivation given for why this language is the basic language for AL is too weak. It is claimed that AL should contain composition, but even the logic of composition from Section 2.4 is not considered as arrow logic. On the other hand, it is claimed that AL has important applications in dynamic semantics (and some related fields), but in the explanations only composition is motivated, while the interpretations for converse and identity are quite doubtful. It is mentioned that one of the possible semantics for AL is the multi-graph one, but it is hardly used in the book. Thus the point of view adopted in this book for AL is too narrow. It is treated mainly as "modal logic for binary relations" as it indeed is and in my view this should be its real name. Arrows are connected with multi-graph intuition and as already mentioned each MDML has such a semantics, so it is natural to consider AL in a general sense and MDML as the same thing.

3. *On the axiomatizations with irreflexivity rules for systems with difference operator.* A novelty in this book is a systematic demonstration of axiomatizations of some "bad" multi-dimensional modal logics, not admitting finite structural axiomatizations, by means of unorthodox derivation systems, based on Gabbay-like irreflexivity rules for difference modality. Normally these are logics based on some full-cube semantics. In this connection, Gargov and Goranko (1993) (see the references in the book) have shown that languages containing difference modality have the same expressive power as the corresponding languages with

universal modality and propositional constants, called names, true at exactly one point. For languages with names, "unorthodox" axiomatizations have also been studied, by means of some special rules for names, invented by Passy and Tinchev (the main reference is their paper cited earlier). So all "unorthodox" axiomatizations in the book can be equivalently given in languages with names and universal modality. The first such axiomatization for some two- and n-dimensional versions of dynamic logic was given by Petkov in 1986 (see his paper cited earlier, and also the Passy and Tinchev paper). Petkov was also the first to consider a two-dimensional tense logic by adding a propositional constant to the language, interpreted as an ordering relation.

4. *On the quality of the text.* The text is very carefully composed in its camera-ready form practically without misprints. I have noticed only a few, which are not serious: page 35, line 9 from below, the second $R_i xy$ should be $R_i xz$; page 195, line 3, the symbol $\leftrightarrow$ should be $\rightarrow$; page 212, footnote 4, the formula $\Diamond\Diamond p \rightarrow p$ should be $\Diamond\Diamond p \rightarrow \Diamond p$.

Dimiter Vakarelov

Department of Mathematical Logic, Faculty of Mathematics and Computer Science, Sofia University, blvd. James Bouchier 5, 1126 Sofia, Bulgaria. dvak@fmi.uni-sofia.bg.

James E. Baumgartner. *Sacks forcing and the total failure of Martin's axiom*. ***Topology and its applications***, vol. 19 (1985), pp. 211–225.

James E. Baumgartner and Peter Dordal. *Adjoining dominating functions*. ***The journal of symbolic logic***, vol. 50 (1985), pp. 94–101.

These papers are both timely and influential additions to the literature and are highly recommended reading (even fifteen years later). This review is not timely, but in requesting that the papers be reviewed at this late date, the editors are likely thinking, as is the reviewer, that many readers would benefit from having these papers highlighted. This is especially so for anyone who is somewhat new to the area of forcing as it relates to properties of the power set of the integers. The quotable results in the papers, especially the second, are interesting consistency results, but in addition the techniques developed are quite innovative and have been very influential in later work.

The main result of the first paper is succinctly stated in the title. The MA number of a poset is the minimum number of dense subsets for which there is no generic filter (i.e., a filter meeting each of the dense sets). The total failure of Martin's axiom is, reasonably, taken to mean that the continuum is larger than $\aleph_1$ and the MA number of every ccc non-atomic poset is $\aleph_1$. The author shows that the model produced by forcing with the so-called side-by-side Sacks (or perfect set) poset (with more than $\aleph_1$ many factors) meets these conditions.

There are about four major and non-trivial techniques that are introduced and well illustrated in this paper. The first is the careful presentation of the side-by-side Sacks poset and the development of its properties. This is a countable-supported product forcing, and the technique of fusion sequences in the usual Sacks poset is generalized to handle the product. This fusion technique is a controlled method of constructing sequences of conditions to have a lower bound that will simultaneously force something about countably many names. In the first instance this shows that the cardinal $\aleph_1$ is preserved, but the technique is more broadly applied than that, both in this and in subsequent papers. In particular it is this technique that is used to show, roughly, that the MA number of a ccc poset will not be increased by the forcing. This is not so direct for a couple of reasons, however, and this argument is the second technique that should be mentioned. Rather than developing a single fusion sequence, the author constructs a binary branching tree of fusion sequences with the result that there are continuum-many incomparable, let us call them, completions of the sequence.

The sequences are constructed in such a way that each is deciding some element of a potential generic subset of a poset, with the result that the existence of the continuum-many incomparable extensions will imply that poset could not have been ccc (and have its MA number increased). The third and fourth techniques involve reflection, factoring, preservation, and names.

Loosely speaking, reflection refers to the technique of associating with an object of indeterminate size, another object of manageable size ($\aleph_1$ in this case) in such a way that taking care of all the manageable-sized objects will result in taking care of all the objects. In proving that Martin's axiom is consistent, one first shows that it is sufficient to consider only posets of cardinality $\aleph_1$. This can be viewed as a (straightforward) reflection result. In fact, one must show, given a poset P, one can associate a *small* poset Q so that the MA number for Q will give information about the MA number for P (each Q is chosen to handle a given collection of dense subsets of P). In forcing terminology, a poset Q is said to be a regular subposet of P if the dense subsets of Q are also dense in P. It is easy to see that the MA number of P is no larger than the MA number for Q in such a situation. Such a Q of size $\aleph_1$ cannot in general be found, but a relativized version is introduced in this paper, that is, relative to a (reflected) inner model. The details are somewhat technical but the ideas introduced and exposited are central to many other applications. Given a ccc poset P in the final Sacks model, a subposet Q (of cardinality $\aleph_1$) is found so that in a model obtained by forcing with an $\aleph_1$-sized subproduct of the Sacks poset, all the dense subsets of Q from that model are dense in P in the final model (but there may well be other dense subsets of Q which are not dense in P). It is this family of $\aleph_1$-many dense subsets of Q that fixes the MA number of Q and P to be $\aleph_1$.

The most interesting result from the paper with Dordal is that forcing with the usual poset for adjoining a dominating real will not *extend towers*. This kind of distinction between different combinatorial problems is crucial in many arguments. The dominating-real forcing *appears* to fill a tower, since it adds a new function from N to N with the property that it is eventually larger than every such function in the ground model. Viewing a function f as a subset of $N \times N$, we can associate with it a subset $U_f = \{(m,k) \mid f(m) \leqq k\}$ of $N \times N$, and notice that in other terminology, the infinite set U_f (for our new function f) is contained, mod finite, in U_g for all old functions g. A family $\mathcal{A}$ of infinite subsets of N is a *tower* if it is a mod finite chain and if there is no infinite set that is mod finite contained in every member of $\mathcal{A}$. It is extremely useful (and difficult) to determine if forcing with a poset will preserve that a family is a tower; one says that the poset *fills* the tower if it introduces an infinite set mod finite contained in every member. Note that the family of U_g's above is already filled by each set $\{n\} \times N$. The length of a tower is the minimal cardinality of a cofinal subset.

This paper also illustrates several of the interesting techniques already mentioned, such as reflection, but in the more familiar context of finite-supported ccc forcing. It is shown that this forcing produces a model in which the continuum can be ω_2, there is no tower of length greater than ω_1, and yet every family of strictly increasing functions from N to N having length ω_2 will in fact be a dominating family. In addition, it is shown that the splitting number $\mathfrak{s}$ in this model is ω_1.

The most interesting—at least to the reviewer—and influential aspect of this paper is its introduction and analysis of a kind of *rank* function on the elements of the poset (the authors call it a derivative). The poset consists of pairs (t, g) where t is a finite increasing sequence of integers (i.e., an approximation to the eventual function f) and g is a ground-model function indicating a promise that f will be larger than g at all integers greater than the length of t. Let *IS* denote the collection of finite increasing sequences of integers. The aspect of this poset that is difficult to analyze is the influence of the uncountably many possible side-conditions g. The rank idea is a technique to extract what we can from the relationship between the members of *IS* and a given name. Given a subset D of *IS*, the derivative D' of D is the set of those $t \in IS$ for which there is a k such that for each integer n there is a $t_n \in D$ that extends t and with k values all larger than n. The idea is that no matter which g is given, there is an n such that (t_n, g) is an extension of (t, g); viz. any n larger than $g(|t| + k)$. It is shown that if D_0 consists of those t for which there is a g such that (t, g) decides the value of a name τ, and inductively $D_\alpha = (\bigcup_{\beta<\alpha} D_\beta)'$, then there is an $\alpha < \omega_1$ such that $D_\alpha = IS$. This gives at least the possibility of proving properties of members of IS by induction on this rank (which is unusual in that the rank decreases as the length of t increases). This technique is cleverly used by considering the set of possible values of τ, i.e., $T_t = \{m \mid (\exists g)(t, g) \Vdash m = \tau\}$. The key observation is that if $t \in D'_\beta$ and $\{t_n \mid n \in N\} \subset D_\beta$ is the sequence witnessing this, then any m that appears in T_{t_n} for infinitely many n, will also be a member of T_t.

ALAN DOW

Department of Mathematics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA. adow@uncc.edu.

JAMES E. BAUMGARTNER. *Bases for Aronszajn trees*. ***Tsukuba journal of mathematics***, vol. 9 (1985), pp. 31–40.

JAMES E. BAUMGARTNER. *Polarized partition relations and almost-disjoint functions*. ***Logic, methodology and philosophy of science VIII, Proceedings of the Eighth International Congress of Logic, Methodology and Philosophy of Science, Moscow, 1987***, edited by Jens Erik Fenstad, Ivan T. Frolov, and Risto Hilpinen, Studies in logic and the foundations of mathematics, vol. 126, North-Holland, Amsterdam etc. 1989, pp. 213–222.

A *base* for an arbitrary tree T of height ω_1 is a collection $\mathcal{B}$ of subtrees of T of height ω_1, such that every subtree S of T includes a member B of $\mathcal{B}$. Here a "subtree" is always assumed to be downwards closed. The problem about the existence of a small base for a given class of trees of height ω_1 is only an instance of a general program to which the author has made substantial contributions. For example, the author's solution of the basis problem for uncountable separable linear orderings is still the finest result in this area.

In the first paper under review, starting from the consistency of the existence of an inaccessible cardinal, the author constructs two models in which the continuum hypothesis is true and false, respectively, and in which every Aronszajn tree (countable levels with no uncountable branches) has a base of the minimal possible size, $\aleph_1$. The use of the consistency of the inaccessible cardinal is necessary in the sense that if every Aronszajn tree has a basis of size $\aleph_1$, then ω_2 is an inaccessible cardinal in the constructible universe. The author also shows that the basis problem for a class of trees of height ω_1 makes sense only if the trees in the class are thin, by showing that if a tree T has height ω_1 and if every node of T has $\aleph_1$ immediate successors then every base for T must have the maximal cardinality $2^{\aleph_1}$. This is done by constructing an "almost disjoint" family of subtrees of T of the maximal possible size. It is interesting that some features of the problem about sizes of almost disjoint families of subsets of ω_1 (another classical problem to which the author has made some of the most important contributions) show up also in the basis problem for subtrees of a given ω_1-tree. For example, adjoining any number of Cohen reals does not increase the basis of a given Aronszajn tree beyond $2^{\aleph_1}$ of the ground model.

The second paper is a contribution to the problem of the existence of a family $\mathcal{F}$ of size κ of pairwise *almost-disjoint* functions from ω_1 into ω, i.e., a family with the property that $\{\xi < \omega_1 : f(\xi) = g(\xi)\}$ is countable for every two distinct members f and g of $\mathcal{F}$. This existence statement can succinctly be written as

$$\binom{\kappa}{\omega_1} \not\longrightarrow \binom{2}{\omega_1}^{1,1}_{\omega}$$

by using the Erdős–Rado arrow notation for the "polarized" partition calculus. The author shows, attributing the result also to Donder and Levinski, that

$$\binom{\omega_2}{\omega_1} \longrightarrow \binom{2}{\omega_1}^{1,1}_{\omega} \quad \text{implies} \quad \binom{\omega_2}{\omega_1} \longrightarrow \binom{n}{\omega_1}^{1,1}_{\omega} \quad \text{for all } n < \omega.$$

Moreover, the author shows for each integer $n \geqq 2$ the consistency of

$$\binom{\omega_3}{\omega_1} \longrightarrow \binom{n}{\omega_1}^{1,1}_{\omega} \quad \text{while} \quad \binom{\omega_3}{\omega_1} \not\longrightarrow \binom{n+1}{\omega_1}^{1,1}_{\omega}.$$

This unravels an interesting discrepancy between the second and the third uncountable cardinals. The discrepancies between ω_2 and ω_3 of this sort are becoming more and more present in modern set theory.

STEVO TODORCEVIC

Université Paris 7, UFR de Mathématiques, 2 place Jussieu, F-75251 Paris cedex 05, France. stevo@logique.jussieu.fr.

GEORGE BOOLOS and RICHARD G. HECK, Jnr. *Die Grundlagen der Arithmetik, §§82–3*. ***The philosophy of mathematics today***, edited by Matthias Schirn, Clarendon Press, Oxford University, Oxford and New York 1998, pp. 407–428.

RICHARD G. HECK, Jnr. *The finite and the infinite in Frege's Grundgesetze der Arithmetik*. Ibid., pp. 429–466.

CRISPIN WRIGHT. *On the harmless impredicativity of $N^=$ ('Hume's principle')*. Ibid., pp. 339–368.

MICHAEL DUMMETT. *Neo-Fregeans: in bad company?* Ibid., pp. 369–387.

CRISPIN WRIGHT. *Response to Dummett*. Ibid., pp. 389–405.

Frege's ***Grundlagen*** (49*5*) §§82–83 attempts to outline a proof that the successor relation, defined in §76 and proved to be one-one in §78, satisfies the condition that every natural number has a successor, thereby completing the sketch of the proof of the principal mathematical contention of the work, namely that the proposition that there are infinitely many natural numbers follows from the principle, variously known as Hume's principle, $N^=$, and the partial contextual definition of the cardinality operator:

> *For any concepts F and G, the number of F's is identical with the number of G's if and only if the F's and the G's are in one-one correspondence.*

Since Frege's "natural sequence of numbers" (as opposed to the sequence of natural—i.e., finite—numbers) includes "Endlos" (which, in ***Grundlagen***, Frege writes ∞_1) or $\aleph_0$, it is necessary to qualify many of the propositions that figure in the proof. Thus, for example, since $\aleph_0$ immediately precedes itself, the theorem (145 of ***Grundgesetze***, 49***16***) that the strong ancestral of the predecessor relation is irreflexive must be formulated as the universal closure of $0P^{*=}n \rightarrow \neg nP^{*}n$, rather than $\neg nP^{*}n$, where P is the relation *immediately precedes*, $P^{*=}$ is its weak ancestral, P^{*} its strong ancestral, and $0P^{*=}n$ is Frege's well-known definition of the natural or finite numbers.

In *Die Grundlagen der Arithmetik, §§82–3*, George Boolos and Richard Heck observe that at a crucial point in the argument of ***Grundlagen***, Frege fails to mark the restriction to finite numbers. This occurs in connection with the proof that the property of immediately preceding the number of one's predecessors is hereditary in the P-sequence, i.e., in the notation of the paper, that $[d : dP\,\#[x : xP^{*}d]]$ is hereditary:

(1) $$dPa \;\&\; dP\,\#[x : xP^{*}d] \;\rightarrow\; aP\,\#[x : xP^{*}a].$$

Grundlagen fails to present a sketch of a correct proof of (1). A natural reconstruction of what is required deploys not (1), but (1′), where (1′) is (1) together with the condition that d be finite. Boolos and Heck point out that in the discussion of ***Grundlagen***, Frege fails to formulate, let alone, establish, the principle of induction that underwrites inferences in which such a restriction to finite numbers figures.

This unsatisfactory situation is resolved in ***Grundgesetze*** where the relevant principle of induction is established (Theorem 152), correct proofs of the subordinate lemmas and principal result (Theorem 155) are given, and where, in the "Analysis" of §114, Frege in effect gives a critical synopsis of the flawed discussion of ***Grundlagen***. Boolos and Heck note that in his "Analysis" Frege expresses the view that (1) is unprovable in his system (i.e., in the second-order theory whose only axiom is Hume's principle), without, however, suggesting that it is false. This is curious, since, given Dedekind's claim (Theorem 159 of $70\frac{1}{2}$***1***, ***Was sind und was sollen die Zahlen?***) that every infinite number is Dedekind infinite, there is an easy argument to (1)—once (1′) has been established. Since the proof of Dedekind's theorem depends on the axiom of choice, Frege's "Analysis" is suggestive of his view of logical methodology, both with respect to what constitutes a provable proposition and with respect to what constitutes a logical truth. The paper concludes with a proof (due to Heck) of (1) that uses only the resources of Frege's system. The paper is a close logical and textual analysis of ***Grundlagen*** §§70–83 together with a comparably close comparison with the emended account of ***Grundgesetze***. There are three useful appendices, one giving collations, another expounding Frege's central theorem, and a third translating into present-day notation §§115, 117, 119, and part of §113 of ***Grundgesetze***.

Heck's *The finite and the infinite in Frege's Grundgesetze der Arithmetik* is an exposition of ***Grundgesetze***'s contributions to the nineteenth-century elucidation of the concepts of finitude and infinity, together with a number of interesting reflections on Frege's work and its connection with that of several of his contemporaries. It continues a series of major contributions Heck has made to the philosophical and logical community by presenting this long-neglected work in a more widely accessible form. Departing from Heck, and using the notation of the paper of Boolos and Heck just reviewed, we have that whenever Q is a functional (many-one) binary relation whose ancestral is irreflexive, Frege says that Q^* is *simple* ("the Q-sequence is simple"). If for every x of the Q-sequence beginning with a there is a y such that xQy, the Q-sequence is said to be *endless*. Theorems 207 and 263 characterize "Frege infinity": G is a *Frege-infinite* concept, in which case the number of G's is Endlos, if and only if the G's can be ordered as a sequence which is both simple and endless:

$$\begin{aligned}\#[x : Gx] = \infty_1 \quad \leftrightarrow \quad & \exists Q[Q \text{ is functional } \& \neg\exists x \,.\, xQ^*x \;\& \\ & \forall x(Gx \to \exists y \,.\, xQy) \;\&\; \exists x\forall y(Gy \leftrightarrow xQ^{*=}y)]\end{aligned}$$

For Frege's notion of finiteness we require the notion, 'x is Q-between a and b,' which we write '$aQ^{*=}xQ^{*=}b$,' and define by the condition:

$$Q \text{ is functional } \& \neg bQ^*b \;\&\; aQ^{*=}x \;\&\; xQ^{*=}b.$$

The major portion of Heck's discussion is organized around Theorem 327, its converse, Theorem 348, and Frege's proofs of them. Together they are expressed, in this notation, by the sentence:

$$0P^{*=}\#[x : Fx] \leftrightarrow \exists Q\exists x\exists y\forall z[Fz \leftrightarrow xQ^{*=}zQ^{*=}y],$$

where, as before, P is the relation *immediately precedes* of Frege's natural sequence of numbers. As Heck says (p. 434), Frege's notion of finiteness is that the number of F's is finite if the F's can be ordered as a simple sequence that *ends*.

Heck's exposition contains a number of remarks on Frege's methodology and choice of proof strategy of which perhaps the most interesting is his conjecture, and the considerations he advances in its favor, that Frege's dissatisfaction with Dedekind's proof of Theorem 159 of ***Was sind … ?*** was prompted by his discovery of its dependence on a formulation of countable choice. Frege's inability to recover this axiom within the system of ***Grundgesetze*** would naturally have forced the question whether it constituted a new logical law or evidence of a "non-logical source of cognition." Either way, Heck argues, it is hard to maintain the view that Frege regarded as nonsensical metalogical questions concerning the status of the laws of logic. While I am in general sympathy with Heck's position, the view he is attacking seems to be something of a strawperson: surely the challenge is to show that (i) Frege exploited a semantic theory, in terms of which the notion of logical truth was to be explained, and by reference to which answers to questions of this sort might be forthcoming, where (ii) the semantic theory is at least in *some* significant sense "model-theoretic." So far as I am aware, no one has gone so far as to maintain that while Frege thought it legitimate to ask what is and is not a law of logic, he refused to entertain questions about the *truth* of logical laws.

Crispin Wright's *On the harmless impredicativity of $N^{=}$ ('Hume's principle')* is an elaboration and defense of the view (first articulated in his book, ***Frege's conception of numbers as objects***, JSL L 252) that Hume's principle constitutes a novel type of "explanation" of the concept of number, what Wright calls "concept introduction via an abstraction principle." I have recently reviewed the main features of this doctrine (JSL LXIII 1598) and so will be brief, devoting most of my attention to Wright's defense of his position here and in his *Response to Dummett*.

Wright is chiefly concerned to address two of Michael Dummett's criticisms. Both criticisms were raised by Dummett in ***Frege, Philosophy of mathematics*** (JSL LXI 1389), but they are articulated again in *Neo-Fregeans: in bad company?*—the article to which the second paper of Wright's is a response. Dummett's *first* objection is that Wright has failed to give the notion of concept introduction via an "abstraction principle" a sufficiently general justification, something that, Dummett argues, is clearly required since it is well known that such principles can lead to inconsistency. To this objection Wright can (and does) reply that it is enough that a procedure that is sometimes flawed is sometimes capable of being shown consistent—as is the case with an abstraction using Hume's principle. Why should it be necessary to show that the procedure will work whatever abstraction principle is employed? One difficulty with this response is that it tends to assimilate concept introduction via an abstraction to the ordinary practice of axiomatization; it thus makes it that much more difficult to understand Wright's insistence that what he is proposing is sharply different from what he calls "the mere axiomatic stipulation of existence." But since it is not immediately obvious what hangs on the contrast with axiomatic stipulation, perhaps this may be put to one side.

Dummett's *second* objection is much more serious: Dummett maintains that Wright is unable to solve the problem to which Frege originally directed his attention, namely: "to secure ... the existence of the objects [of number theory]" (p. 369) and "to show us how to attain the conception of [a countably infinite] domain" (p. 381) without relying on an appeal to intuition or the facts of experience. This second objection calls into question the philosophical interest of Wright's program. (The main substance of this objection precedes Wright's book, and can be found, for example, in ***Frege, Philosophy of language***, Duckworth and Harvard University Press, 1973, where it is urged against Frege; see especially pp. 532ff. For Wright's response to this formulation, see the long note to Section xvi of ***Frege's conception of numbers as objects***, pp. 180–184.)

It is generally believed that Frege achieved neither of the goals Dummett attributes to him. What is unusual about Dummett's assessment is his claim that Frege's failure had little to do with Russell's discovery of the contradiction. Dummett of course grants that Russell's paradox affected Frege's theory of classes; but the sole purpose of *that* theory was to secure the logical character of Hume's principle by facilitating its derivation from the explicit definition of the cardinal numbers as equivalence classes of concepts under one-one correspondence. According to Dummett, the contradiction, serious though it was for Frege's concept of class and the subsequent development of a theory of the *real* numbers, has only obscured the basic problem with Frege's whole approach. The problem is clarified if we grant Frege the logical character of Hume's principle and then ask whether his use of it, in combination with the "context principle" (*Only in the context of a sentence does a word have reference* [Bedeutung]), actually succeeds in addressing the issues he sought to resolve.

Frege's idea appears to have been that numerical singular terms are referential—and thus, the existence of the objects of number theory secured—because with Hume's principle we are in possession of a clear criterion by which we can say when the same number has been "given to us" in two different ways, as the number of one or another concept. Our "access" to the numbers is thus mediated by our recognition of the truth of Hume's principle, and the same principle also serves as the only substantive premiss in the proof of the infinity of the numbers. Since the same methodology underlies Frege's introduction of classes—with, of course, the difference that there the relevant contextual definition (Basic Law V) is inconsistent—we would do well to see whether Frege's procedure is defensible *at all*. This is the rationale for Dummett's inquiry into the success of Frege's account of our reference to the numbers in terms of the context principle and a *consistent* contextual definition (Hume's principle), an inquiry which leads Dummett to conclude that there is something fundamentally misguided about this attempt of Frege's to secure the existence of the objects of number theory and account for our conception of a countably infinite domain. This contrasts with Wright who begins from the observation that Frege's proof of the infinity of the number sequence depends only on Hume's principle and, in particular, does *not* require Frege's inconsistent theory of classes. For Wright, serious though the contradiction was, it has obscured one of Frege's fundamental *achievements*. The question therefore naturally arises as to how there can be such divergent opinions, given what appears to be overriding agreement on the role of classes, the significance of the inconsistency of Frege's theory, and the assessment of the mathematical achievement of ***Grundlagen*** (modulo the slip discussed earlier in connection with the paper of Boolos and Heck) and ***Grundgesetze***.

Dummett grants that we can attain the conception of the infinity of the number sequence from Frege's proof of his theorem but questions whether we may infer from this the conception of a countably infinite *domain*. This latter question is, for Dummett, necessarily linked to the question whether and in what sense Frege secured the existence of the objects of number theory. While Frege's theorem shows the numbers to form an infinite "totality," this falls short of securing the existence of what Dummett calls a "*definite* totality" or domain. But unless the numbers are shown to constitute a domain, we have not justified the view that all our statements about them have classical truth conditions, and this leaves our claim to have secured their existence open to interpretation; at best we have achieved a conception according to which every sentence we might formulate about the numbers has determinate *justification conditions*. According to Dummett, Frege relied on a procedure that was capable at most of securing justification conditions for the sentences of number theory, while all along supposing himself to have secured classical truth conditions for them. On the assumption that Hume's principle is one we are justified in accepting, the proof shows how, given any natural number n, we are always justified in accepting the sentence that asserts the existence of a successor to n: indeed, it gives us the successor as the number of n's predecessors and thus shows how the concept of any finite sequence of numbers can always be extended to a more inclusive one. But such a provision of justification conditions is compatible with a denial of what Dummett has called the "external" reference of numerical singular terms in a way in which properly realist truth conditions are not. (Unless I have misread him, the terminology of "domains" versus "totalities" is simply short-hand for the contrast between providing classical truth conditions—appropriate to "domains"—as opposed to justification conditions—appropriate to "indefinite totalities." In particular, the question whether it makes sense to quantify over "absolutely everything," whether domains of quantification must always be sets, and other related matters, though subjects concerning which Dummett has opinions, are not essential to the question at issue.)

For Dummett, the mistake of Frege—and of Wright, who follows Frege on this point—is to suppose that Hume's principle can both have classically determinate truth conditions and serve to "introduce" the numbers. The notion of classically determinate truth conditions presupposes that we have independent access to whatever our statements are held to be *true of*, a presupposition that is violated by the "impredicativity" of Hume's principle, that is, by the fact that the numbers introduced on the left occur within the range of the variables bounded by the first-order quantifiers occurring on the right (in the explicit definition of one-one correspondence). Dummett argues that this combination of the principle's impredicativity and its deployment as a kind of partial contextual definition gives rise to a vicious circularity in Frege's and Wright's methodology, one that also affects the interpretation of the cardinality operator. Since for Frege the definiteness of a concept requires that it be determinate for every object whether or not it falls under the concept, if the domain of objects is not fixed, neither is the domain of concepts; nor, therefore, is any function from concepts to objects. To address the circularity, we must give up the idea that Hume's principle and all those sentences whose proofs depend upon it have been given anything more than justification conditions. Frege can still be represented as having established that his conceptual analysis of numerical identity yields a statement which implies that *natural number* is a Dedekind infinite *concept*. But there can be no pretense of his having secured or guaranteed the existence of a *domain* of objects for number theory; and the task of providing an account of the basis for our initial assumption, that the numbers are available for inclusion in the domain of the individual variables of Hume's principle, remains uncompleted. The difficulties Dummett is alleging would not have arisen if Hume's principle were not impredicative (see e.g. the passages quoted by Wright in footnote 25 of pp. 349–350 and on p. 351). But if it were not impredicative, it would be incapable of supporting the proof of the infinity of the number sequence, since the proof turns on the possibility of forming, for each n, a numerical concept under which fall just the first n numbers. (Although on its intended interpretation Hume's principle quantifies over *all* concepts of divided reference, it would suffice if all numerical concepts were included within the range of the second-order variables. But even under this weakening, the successful use of the principle in Frege's proof requires that the numbers fall within the range of the first-order variables, since they are needed to form the numerical concepts on which the success of the argument depends.)

Is the impredicativity of Hume's principle indispensable to its fulfilling the role of *introducing* the numbers by an abstraction principle? In other words, leaving aside its role in the proof of the infinity of the number sequence, might it be the case that Hume's principle could, in Wright's sense, "introduce" the number concept if the numbers did *not* lie within the scope of the first-order variables occurring in the explicit definition of one-one correspondence? Wright's answer to this (latter) question is "No" (see pp. 351 and 356). The reason is that our concept of number is one of extreme generality: anything can be counted, including, in particular, the numbers themselves. Hence, if Hume's principle is to serve to introduce a concept of *number*, the numbers must fall within the range of the first-order variables, so that concepts under which the numbers themselves fall will also lie within the domain of the cardinality operator. There is a sense, therefore, in which Wright is prepared to make his position more difficult by requiring the impredicativity of Hume's principle, both at the level of its role in the proof of Frege's theorem *and* in its role as a partial contextual definition.

Wright's response to the charge of circularity is to argue that since impredicativity is not a barrier to our *understanding* of the partial contextual definition, it can serve as an effective "explanation" of the concept of number. The paper contains an extended argument to this effect and the *Response* argues that a lacuna Dummett points to in this argument can be closed. Without entering into the detail of Wright's discussion, let me try to clarify what it purports to establish. Assume, to begin with, that we may ignore the so-called "Julius Caesar problem"—suppose, in other words, that if the first-order variables range over a proper "domain" we may assume that the numbers can be distinguished within that domain. Indeed, if we ignore the Caesar problem, we may assume that the domain consists of *just* the objects constituting Frege's natural sequence of numbers. The task Wright sets himself is that of showing how someone might come to a conception of this domain given an understanding of the logical resources of a higher-order language. The problem Wright addresses is therefore to show how, under these assumptions, one can advance from a grasp of "simple" sentences containing the cardinality operator to an understanding of its occurrence throughout the proof of Frege's theorem, where, roughly speaking, the notion of simplicity is that simple sentences apply the cardinality operator only to those predicates whose understanding requires no arithmetical competence. (On the assumption that the domain consists only of numbers, $x \neq x$ and its truth-functional equivalents are, perhaps, the only predicates that are simple in this sense.) Let us grant that should Wright succeed in showing this, it is a short step to the conclusion that he has explained our grasp of the proof as a whole, and thus, given the content of the theorem, that he has explained our grasp of a Dedekind infinite concept. Dummett, of course, will argue that this is insufficient and that Wright needs to show that his argument has secured the conception of a Dedekind infinite *domain*. It is unclear to me how this dispute between Wright and Dummett might be resolved.

The controversy between Dummett and Wright brings to the surface—but does not resolve—another difficulty with Wright's articulation of the neo-Fregean position. Recall Frege's idea that we can secure the thesis that numerical singular terms are referential because we are in possession of a criterion of identity for them. This account of our access to the numbers relies fundamentally on our recognition of the truth of Hume's principle. What we require therefore is an account of the sense in which that principle is true—and an account of how it is known to be true—since this is what is needed to infer, via the context principle, that numerical singular terms refer. Wright's view of this is not altogether straightforward. For example, any suggestion that Hume's principle is an *analytic* truth, *analytic of* the cardinality operator, must be understood in light of Wright's "official" position, according to which the principle is freely laid down as a *stipulation* (cf. especially p. 341). But this only raises a more basic issue, namely that of understanding how it is possible to secure the reality of the numbers—or, for that matter, of any domain of objects—by a mere stipulation. Wright maintains that the representation of *Hume's principle* as a kind of stipulation does not prejudice the question whether *the existence of the numbers* is a matter of stipulation. Wright holds that the existence of the numbers is something discovered rather than stipulated—it follows from an "explanation" rather than "a mere axiomatic stipulation"—even though our knowledge of their (necessary) existence is ultimately derived from Hume's principle, a principle whose truth *is* a matter of stipulation. This account might be credible as a view of conventional *explicit* definitions: it is precisely because ordinary definitions are not creative that we can say that anything established or "discovered" on their basis does not really depend on them, and, in consequence, does not share their epistemic basis. But it can hardly be maintained in the case of a contextual definition like Hume's principle which, even if a stipulation, is not a stipulation of the same character as an explicit definition, since it is creative over the theory to which it is added, namely, pure second-order logic. This is where the analogy between "concept introduction via an abstraction principle" and the methodology of concept introduction by ordinary explicit definition breaks down. To my mind, Wright does not adequately address this issue. Instead he has chosen to concentrate on the question whether the surface grammar of the left-hand side of Hume's principle can be taken seriously, whether, in other words, the apparent deployment of numerical singular terms *as* singular terms is justified. As Wright puts it, one has "to read the left-hand sides of the appropriate [abstraction] principles not merely as notational variants of the right-hand sides, but in a way which is constrained by their surface syntax" (p. 404). Once this is established, Wright argues, no further question can arise over whether such terms genuinely refer, and there is "no good sense in which their reference could be stigmatized as 'semantically idle' " (p. 404). But the difficulty, as I see it, is with the transition from truth-by-stipulation to truth, since it is clearly the latter notion that the use of the context principle requires. This transition is deeply perplexing and remains the chief stumbling-block to accepting Wright's claim to have secured the reference of numerical singular terms.

William Demopoulos

Department of Philosophy, Talbot College, University of Western Ontario, London, Ontario N6A 3K7, Canada. wgdemo@julian.uwo.ca.